BIOCELL
2020 44(4): 777-803

Tech Science Press

Identification and analysis of AP2/ERF gene family in tomato under abiotic stress

Xiaolin ZHU[1,3]; Xiaohong WEI[1,3,*]; Baoqiang WANG[2,3]; Xian WANG[2]; Mingjun ZHANG[4]

[1] College of Agronomy, Gansu Agricultural University, Lanzhou, 730070, China
[2] College of Life Science and Technology, Gansu Agricultural University, Lanzhou, 730070, China
[3] Gansu Provincial Key Laboratory of Aridland Crop Science, Lanzhou, 730070, China
[4] Gansu Agricultural University of Plant Protection College, Lanzhou, 730070, China

Key words: AP2/ERF gene family, Tomato, Abiotic stress, Expression analysis

Abstract: AP2/ERE-type transcription factors, as a type of plant-specific transcription factors, play a key role in plant biotic and abiotic stress. Meanwhile, they have been studied in many plants, but rarely in tomatoes. In this study, we performed a genome-wide analysis of the SlAP2/ERF gene family of tomato, and finally identified 29 SlAP2/ERF genes and divided them into different subfamilies. At the same time, its basic physical and chemical properties were analyzed. We also constructed phylogenetic trees with 30 *Arabidopsis* AP2/ERF proteins and 28 potatoes AP2/ERF proteins to ensure conservative homology between them. In addition, we mapped 29 SlAP2/ERF transcription factors on 10 different chromosomes; and identified 43 responsive plant hormones, responsive light signals, tissue-specific expression and stress response elements from 2000bp upstream of the promoter region, and we analyzed conserved motifs and gene structures of SlAP2/ERF. The tertiary structure of SlAP2/ERF protein was constructed by homology modeling, and the protein-protein interaction network was constructed based on *Arabidopsis* Thaliana. Finally, the expression pattern of tomato in different tissues was studied by using gene expression database, and the expression level of tomato under abiotic stress was detected by q-RT-PCR. These results provide comprehensive information for further study of the function of the SlAP2/ERF gene family.

Introduction

In most cases, plants are subjected to different biotic and abiotic stress, such as drought, salinity, and low temperature, which seriously affect their yield and quality. Therefore, during the long-term and continuous development, plants have evolved various defense mechanisms to resist various stresses, including carbohydrate accumulation, increased antioxidant enzyme activity, and induced resistance gene expression (Chinnusamy *et al.*, 2004; Hobert, 2008). Transcription factors are proteins that specifically interact with cis-regulatory element in the promoter region of eukaryotic genes and can activate or inhibit the transcription initiation of other genes, and it can regulate the gene expression by interacting with other transcription factors (Singh *et al.*, 2002). There are a lot of transcription factors in plants, which play an important role in the process of plant growth, physiological and biochemical reaction, signal transduction and response to environment. Studies have shown that overexpression of transcription factor genes in response to stress can promote the expression of genes associated with stress resistance downstream and improve the stress resistance of genetically modified crops (Liu *et al.*, 1998; Gilmour *et al.*, 2000; Sakuma *et al.*, 2006). There are 5.9% of the total genes are transcriptional factors in *Arabidopsis* (Thamilarasan *et al.*, 2014).

Transcription factors are divided into AP2/ERF, WRKY, B-zip, NAC and MYB families according to their DNA binding domains (Thamilarasan *et al.*, 2014; Zhang *et al.*, 2010). AP2/ERE-like transcription factors are plant-specific transcription factors, which are characterized by at least one AP2 binding domain. The domain contains 60 to 70 amino acids and is highly conserved. Meanwhile, AP2/ERE transcription factors are involved in plant development and stress pathway (Gutterson *et al.* 2004; Licausi *et al.*, 2013). According to the AP2 domain, the AP2/ERE family is divided into three subfamilies. Among them, the AP2 subfamily contains two AP2 domains, the RAV subfamily

*Address correspondence to: Xiaohong Wei, weixh@gsau.edu.cn
Received: 14 February 2020; Accepted: 18 May 2020

Doi: 10.32604/biocell.2020.010153

includes one AP2 and one B3 domain, and the ERF subfamily contains one AP2 domain (Sakuma *et al.*, 2002). The first AP2/ERF gene, *APETALA2*, was found in *Arabidopsis thaliana*, and its function is mainly related to flower development (Jofuku *et al.*, 1994). The results show that AP2/ERF family is a plant-specific TF family, and its subfamily plays a key role in regulating plant growth and development, such as regulating flower, fruit development and leaf shape (Boutilier *et al.*, 2002; Aya *et al.*, 2014). RAV subfamily may be a negative regulator in plant growth and development, and it plays a key role biotic and abiotic stress in plant (Hu *et al.*, 2004; Sohn *et al.*, 2006; Licausi *et al.*, 2013). The ERF subfamily, which is mainly responsible for plant responses to environmental and biological stresses, plays a key role in plant resistance to stresses, and can activate the expression of stress-resistant genes, thus increasing plant resistance to stresses (Liu *et al.*, 2013). The TF gene of ERF-A1 group has been cloned from several plants, such as *Arabidopsis thaliana* (Novillo *et al.*, 2012), wheat (Soltész *et al.*, 2013) and Tea (Wu *et al.*, 2015). In addition, the role of these genes in enhancing cold resistance of plants has also been confirmed. In maize, Yeast one-hybrid test showed that *ZmERF1A* gene was involved in the expression of cold response genes (Hu *et al.*, 2004). Overexpression of *lsERF2a* increased the salt tolerance of genetically modified crops (Kudo *et al.*, 2014). ERF is mainly related to low temperature, disease resistance, hormones, and high salt (Singh *et al.*, 2002). AP2/ERF transcription factors have been shown to play a key role in flower development in *Arabidopsis* (Kunst *et al.*, 1989), some members of the AP2/ERF family are also link factors in various adversity signaling pathways (Liu *et al.*, 2013), and AP2/ERF genes have been identified in many plants, including *Arabidopsis* (Nakano *et al.*, 2006), Rice (Rashid *et al.*, 2012), Cauliflower (Li *et al.*, 2017), Cabbage (Song *et al.*, 2013), and Pear (Li *et al.*, 2018).

Tomato (*Solanum Lycopersicum*) is not only an important economic vegetable, but also an important model plant for studying plant pathogen interaction (Jiang *et al.*, 2019). In 2012, the world's annual tomato production exceeded 160 million tons, where China accounted for 50 million tons (http://faostat.fao.org). The tomato genome has been sequenced and assembled by the International Tomato Genome Sequencing Project (http://solgenomics.net/Organism/Solanum Lycopersicum/genome). Simultaneously, tomato is economically important and is a model species for studying fruit maturation (Consortium, 2012). So, a high-quality genome sequence and more than 30,000 kinds of proteins have been obtained in tomato. In previous tomato AP2/ERF studies have shown that, over-expression of tomato *JERF1* in tobacco can improve salt tolerance and low temperature resistance (Wu *et al.*, 2007), and over-expression of tomato *JERF3* increases tolerance to drought, low temperature, high salt, and osmotic pressure. Overexpression of tomato *TERF2/LeERF2* in tomato and tobacco enhances tolerance to freezing damage (Zhang *et al.*, 2010), and overexpression of tomato *TSRF1* in rice can increase its tolerance to high osmotic pressure and drought (Quan *et al.*, 2010). Overexpression of tomato *SlERF3ARD* promotes the expression of PR genes such as *PR1, PR2*, and *PR5*, and significantly increases resistance to bacterial wilt disease (Pan *et al.*, 2010). Li *et al.* (Li *et al.*, 2011) found that the overexpression of *AtCBF1* transgenic lines in tomato, through continuous expression of RAV transcription factors, ERF family genes, and disease-related genes (PR) significantly increased resistance to bacterial wilt. Further study showed that *SlERF5* could regulate the expression of *AtCBF1* and *PR* genes by regulating the expression of SlRAV gene, which leads to increased resistance to disease. Overexpression of tomato *Pti4* gene in *Arabidopsis* thaliana could improve the resistance to powdery mildew and bacterial spot disease (Venkatesh *et al.*, 2013). These studies suggest that SlAP2/ERF transcription factors play a key role in both biotic and abiotic processes.

However, tomato grown in northwest China is susceptible to various abiotic stresses such as drought and freezing, and its stress resistance is poor, while AP2/ERF transcription factors play a key role in plant response to adversity and it is also essential for the normal development of many plants and interacts with various proteins to regulate gene expression. By given the importance of the AP2/ERF family, therefore, this study is based on the database of tomato genome and transcriptome. Twenty-nine AP2/ERF transcription factors were identified and systematically analyzed in tomato by bioinformatics, including the conserved domains and motifs, phylogeny, basic physical and chemical properties, gene structure, promoters and their expression in different tissues. At the same time, the expression pattern of AP2/ERF gene under abiotic stress was studied, which laid a solid foundation for the application of AP2/ERF transcription factor in genetic improvement of tomato.

Materials and Methods

Plant materials and treatments

Identification of AP2/ERF transcription factor family genes in tomato

Tomato genome data from the Phytomoze (Version 12.1) database (https://phytozome.jgi.doe.gov/pz/portal.html) was downloaded. Also downloaded the Hidden Markov model file (PF00847) (Finn *et al.*, 2008) of AP2/ERF transcription factor in the Pfam database (http://pfam.xfam.org/). Using Hmmer 3.0 software based on hidden Markov Model (HMMs), the AP2/ERF gene family of tomato was searched and identified by $e \leq 0.01$. The identified genes were submitted to Pfam (http://pfam.xfam.org/family), NCBI-CDD (https://www.ncbi.nlm.nih.gov/cdd/) and SMART (http://smart.embl-heidelberg.de/) online tools to predict the conserved domain of AP2/ERF protein, recorded the location of the conserved domain, and to eliminate the false-positive gene sequences (Letunic *et al.*, 2015).

Analysis of physical and chemical properties of Tomato AP2/ERF family

The Amino acid number, isoelectric point, molecular weight and hydrophobicity of 29 tomato AP2/ERF candidate proteins were analyzed in Expasy (https://web.expasy.org/protparam/).

Evolutionary analysis of AP2/ERF transcription factor family in tomato

28 AP2/ERF transcription factors of potato and 30 AP2/ERF transcription factors in *Arabidopsis* were identified from the Phytomoze database. Using the Clustal W function in MEGAX software, we compared the conserved amino acid sequences of identified AP2/ERF family proteins in tomato, *Arabidopsis thaliana*, and potato. The results were used to construct the phylogenetic tree using the proximity method and the Poisson model, the number of repetitions is 1000 (Song *et al*., 2014) and the other parameters were set to default. The tree of evolution is beautified through the online software named Evolview (http://120.202.110.254:8280/evolview).

Structure analysis and chromosomal location of AP2/ERF transcription factor gene in tomato

The Gene Structure Display Server GSDS tool (http://gsds.cbi.pku.edu.cn/) (Hu *et al*., 2015) was used to identify exon and intron structure of the gene. The conserved amino acid motifs were analyzed (Bailey *et al*., 2006) by using MEME online software (http://meme-suite.org/tools/meme). The number of motif searches was set to 10, and other parameters were default. According to AP2/ERF gene's position in the tomato genome, 29 AP2/ERF genes were located on corresponding chromosomes by Mapchart software (Voorrips, 2002).

Promoter analysis, homology modeling and protein interaction network mapping of tomato AP2/ERF family genes

The upstream 2000 bp sequence of the SlAP2/ERF family gene was extracted from the tomato genome database as a promoter region, and the gene cis-regulatory element was analyzed by Plantcare online software (Lescot *et al*., 2002). The SlAP2/ERF protein's three-level structure was modeled by Swiss-model server (Kelley *et al*. 2015). We used String (http://Stringdb.org/) online software (Franceschini *et al*., 2013), based on *Arabidopsis*, to predict protein-protein interaction networks based on tomato and *Arabidopsis* homologous genes.

Analysis of the gene ontology (GO)

The function of SlAP2/ERF family in tomato was annotated by using Blast2GO program (Ye *et al*., 2006) (http://www.geneontology.org/) with default setting apart from e-value: 1.0 e^{-3}.

Gene expression analysis

Database of tomato gene expression (Tomato Functional Genomics Database) was used to search the expression data of SlAP2/ERF gene in different tissues. FPKM (Fragments Per Kilobase of transcript per Million Fragments mapped) as an indicator of gene expression level. Heat maps were created by using TBtools.

The total RNA of each sample (Tomato Leaf of different treatments, including 20% PEG, 100 mmol/L Nacl, 100 umol/L ABA, 100 umol/L SA, 4°C low temperature) was extracted with Rneasy small plant reagent box (Qiagen), and the cDNA was prepared with Superscript™ III reverse transcriptase reagent box (Invitrogen). Real-time fluorescent quantitative pcr primers are shown in the following Tab. 1 and have been synthesized commercially. qRT-PCR analysis was performed in abi-viia 7 real-time PCR systems of American applied biosystems by 2 quantities ect-sybr-green-pcr-mix (Qiagen). Tomato actin (Action) gene was used as endogenous control. The relative gene expression level was calculated by $2^{-\Delta\Delta Ct}$ (Livak and Schmittgen, 2001). Each experiment was repeated three times with a separate RNA sample.

Results

Identification of AP2/ERF gene family members and analysis of physicochemical properties

The database of tomato proteins was searched by HMMM 3.0 software, and the conserved domains were identified by PFAM, SMART and NCBI-CDD. A total of 29 genes with typical AP2 domains were identified (Fig. 1). The analysis of the basic physical and chemical properties of AP2/ERF gene family (Tab. 2) showed that the number of amino acids were 120–677, the relative molecular weight was 13755.10-75240.22, there were 14 basic proteins (PI > 7) and 15 acidic proteins (PI < 7). An interesting phenomenon is that the hydrophobic index of all proteins in this family is less than 0, and they are all hydrophilic proteins.

TABLE 1

Primer information of 10 SlAP2/ERF genes by qRT-PCR

Gene name	Forward primer	Reverse primer
SlAP2-02	TGGACCAAGCTATAACACGCCATC	AAGCAGAAGCCGAAGCAGAAGAAG
SlAP2-05	GTTCGGAAGAGGAAACGAGGTGAG	TGTGGCGGGCAGTAGAAAATTCG
SlAP2-07	AACGGAGACACTTGGCAACACG	GTCCACGGCGGCTCTTCTTAATAG
SlAP2-08	GAGGAGGTATGGCTCACAGTTTGC	TGTTGCTGCTGCTGATGTCTAGTG
SlAP2-09	AGAGGGGTAAGAAGGAGGCCATG	AAGCCGCCTTGTCATAAGCCATAG
SlAP2-14	ACAACCAGCAGCAGCAGATGAC	GATAGCAGAGACGGGTGGCATTG
SlAP2-15	GAAGGAGGCCATGGGGTAAATA	CAATGCAGCTTCTTCAGCAGTGTC
SlAP2-22	GGAGAGAAGCGGTGGAAACTGAAG	TGACGGCGTTAATGGACATGACTC
SlAP2-24	CTTACGATGCCGCTGCCAGAG	TCCACGGTACTGCTCCCACAC
SlAP2-26	ATCGGAACCGCCACTTGTCAATG	ATTCCCACCGGCCAGTCCTC

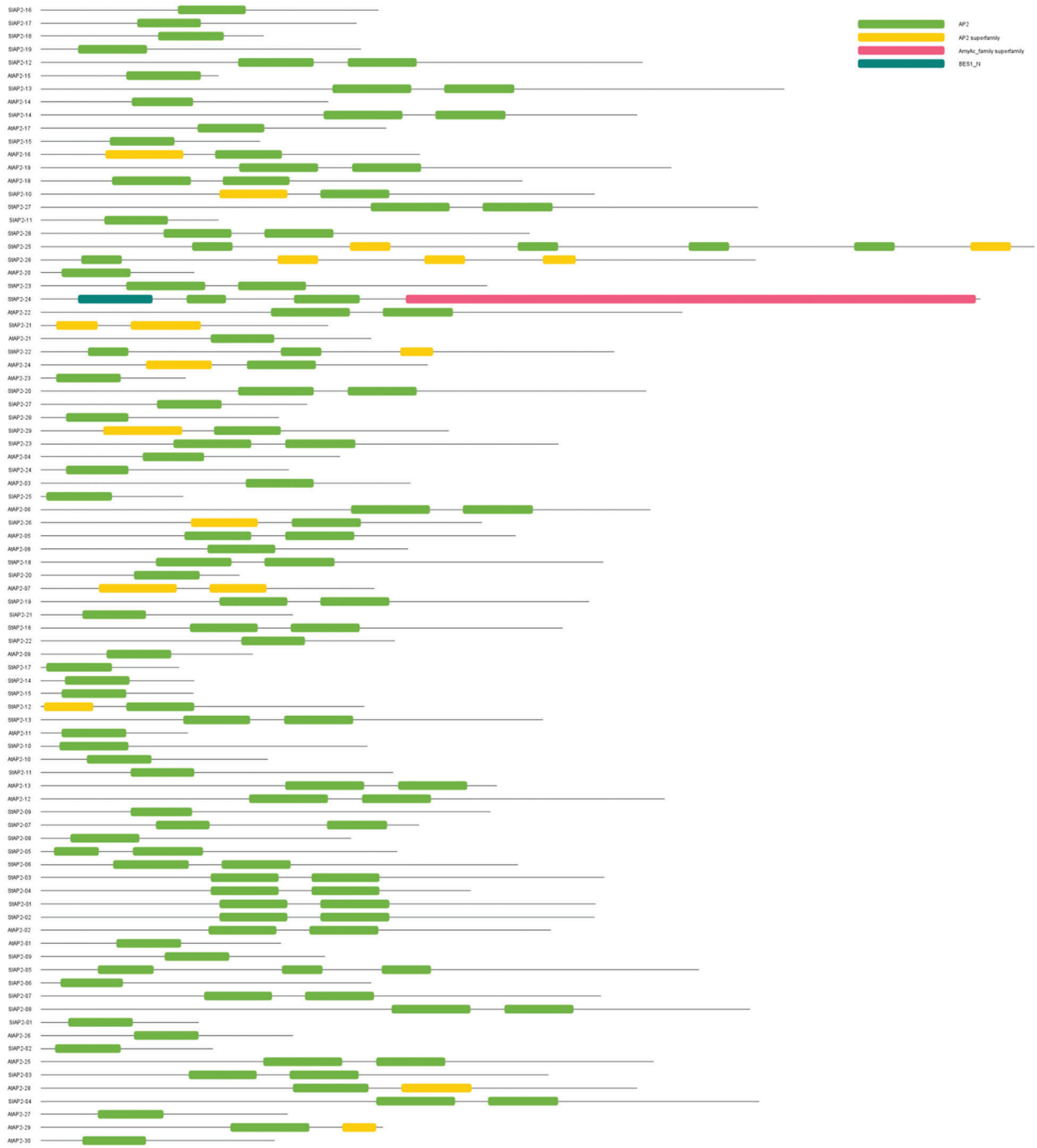

FIGURE 1. AP2 conserved domain of SlAP2/ERF proteins in tomato.

Phylogenetic analysis of members of AP2/ERF gene family

In order to study the phylogenetic relationship of tomato, we constructed a phylogenetic tree of 87 AP2/ERF transcription factors amino acid sequences of tomato, potato and *Arabidopsis thaliana* by using MEGA software based on the ClustalW multiple sequences (Supplementary File Tab. S1). A total of 8 branches were identified in the phylogenetic tree (Fig. 2), each group containing a different percentage of gene members. In the first group, there were only 3 SlAP2/ERFs and 4 StAP2/ERFs; in the second group, 9 SlAP2/ERFs, 3 StAP2/ERFs and 4 AtAP2/ERFs; in the third group, 1 SlAP2/ERF and 4 AtAP2/ERFs; in the fourth group, 1 SlAP2/ERF, 3 StAP2/ERFs and 7 AtAP2/ERFs; and in the fifth group, 2 SlAP2/ERFs and 2 StAP2/ERFs; group 6 has 1 StAP2/ERF and 1 AtAP2/ERF; group 7 has 4 SlAP2/ERFs, 8 StAP2/ERFs and 3 AtAP2/ERFs; and group 8 has 7 SlAP2/ERFs, 8 StAP2/ERFs and 11 AtAP2/ERFs. The phylogenetic tree showed that there were intraspecific and intraspecific homology between tomato, *Arabidopsis*, and potato. Five pairs of orthologous homologous genes between *Arabidopsis* and tomato, and 10 pairs of orthologous homologous genes between tomato and potato. There were 5 pairs of paralogous homologous genes between *Arabidopsis* thaliana and potato. There were 2 pairs of paralogous homologous gene pairs in tomato, 5 pairs of paralogous homologous gene pairs in potato and 5 pairs of paralogous homologous gene pairs in *Arabidopsis*. Meanwhile, the results showed that all the genes in the seventh and eighth branches contained two

TABLE 2

Basic characteristics of tomato (AP2) genes

Phytozome gene ID	Gene name	Protein length (aa)	Molecular weight (Da)	pI	GRAVY	Chr.position
Solyc09g089910.1.1	SlAP2-01	143	16158.76	5.77	−0.891	Chr09:699945342-6999495
Solyc09g066350.1.1	SlAP2-02	156	17764.37	7.08	−1.104	Chr09:65014582 -65015052
Solyc10g084340.2.1	SlAP2-03	462	51081.25	6.06	−0.677	Chr10:64051146-64055505
Solyc04g077490.3.1	SlAP2-04	654	71838.75	7.16	−0.646	Chr04:62512499 -62519587
Solyc04g009450.1.1	SlAP2-05	599	67226.97	9.11	−0.790	Chr04: 2859935 -2861734
Solyc02g067020.1.1	SlAP2-06	300	34110.17	5.60	−0.692	Chr02:37828480-37829382
Solyc02g093150.3.1	SlAP2-07	510	56999.91	6.69	−0.803	Chr02:54715052 -54718090
Solyc02g092050.3.1	SlAP2-08	646	72002.33	6.62	−0.806	Chr02:53895704 -53902352
Solyc02g077360.1.1	SlAP2-09	258	30015.52	4.96	−0.791	Chr02:42885146 -42885922
Solyc02g064960.3.1	SlAP2-10	504	57299.16	8.51	−1.018	Chr02:36641329 -36645265
Solyc02g077370.1.1	SlAP2-11	161	18051.27	7.69	−0.526	Chr02:42890880 -42891365
Solyc11g061750.2.1	SlAP2-12	548	60739.39	6.10	−0.816	Chr11:48672010-48678655
Solyc11g008560.2.1	SlAP2-13	677	75240.22	6.27	-0.818-	Chr11:2752880- 2752962
Solyc11g010710.2.1	SlAP2-14	543	59744.62	6.13	−0.628	Chr11:3760004- 3760583
Solyc11g006050.1.1	SlAP2-15	199	22541.83	4.83	−0.729	Chr11: 854696 -855295
Solyc12g056980.1.1	SlAP2-16	307	33663.72	8.65	−0.539	Chr12:64012226 -64013239
Solyc12g042210.2.1	SlAP2-17	287	33337.47	6.90	−1.108	Chr12:58194225-58194544
Solyc01g090300.2.1	SlAP2-18	202	22735.46	7.80	−0.651	Chr12:58194225-83839843
Solyc01g091760.2.1	SlAP2-19	291	32393.42	7.10	−0.801	Chr01:83839127 -85217581
Solyc01g090310.2.1	SlAP2-20	180	20417.70	7.88	−0.818	Chr01:83847469-83848168
Solyc08g007830.1.1	SlAP2-21	299	26020.79	4.99	−0.767	Chr08:2339145 -2339834
Solyc08g078190.1.1	SlAP2-22	322	36263.72	6.10	−0.565	Chr08:62147117 -62148085
Solyc07g018290.3.1	SlAP2-23	471	52749.14	8.34	−0.806	Chr07:10014430 -10019198
Solyc07g053740.1.1	SlAP2-24	225	24087.23	9.30	−0.434	Chr07:62294729 -62295406
Solyc03g005500.1.1	SlAP2-25	120	**13755.10**	6.07	−0.893	Chr03: 383587 -383976
Solyc03g044300.3.1	SlAP2-26	401	44944.48	7.68	−0.864	Chr03:8790516 -8795225
Solyc03g093610.1.1	SlAP2-27	242	26760.38	7.68	−0.433	Chr03:56458291 -56459019
Solyc03g006320.1.1	SlAP2-28	216	22442.29	9.04	−0.303	Chr03: 924296 -924946
Solyc03g117720.3.1	SlAP2-29	371	42093.77	6.77	−0.893	Chr03:68259122 -68263665

Note: GRAVY represents grand average of hydropathicity.

typical AP2 domains, belonging to the AP2 subfamily. All the genes in the second branch to the sixth branch contained an AP2 domain, belonging to a typical ERF subfamily. The SlAP2-05, StAP2-22, StAP2-25, and StAP2-26 genes in the first branch all contain more than 2 AP2 domains.

Gene structure analysis and chromosome localization of tomato AP2/ERF transcription factor

Since the global pattern of intron positions in some gene families served for the evolution of typical mark (Del Campo *et al.* 2013; Rogozin *et al.*, 2005), and the absence of introns could lead to the diversity of genetic structure, so we investigated SlAP2/ERF family introns-explicit substructure, to further study the evolution of their relationship. The results (Fig. 3B) showed that all 16 genes of class I and III of the SlAP2/ERF family have not contain intron structure, the 11 genes in the II groups all contain introns, and the number is between 7 to 9, and the main gene structure is 9 exons and 8 introns (8 of 11 SlAP2s). In order to further study the structural diversity of tomato AP2/ERF protein, ten conservative motifs were revealed by Meme (Fig. 3C). We can conclude that AP2/ERF proteins clustered in the same group have similar motif domains (Fig. 3A), and all 29 genes contain motif1 and motif2, in which motif1 mainly contains the conserved domain L3-G5-A12 (Leu3-Gly5-Ala12) and motif2 contains a highly conserved domain Y1-R2-G3 (Tyr1-Arg2-Gly3). A gene with an AP2 domain, it can be divided into ERF subfamily and DERB subfamily according to the amino acid category of 14 and 19 (14 is valine, 19 is glutamate belongs to DERB subfamily, 14 is alanine, 19 is L-aspartic acid belongs to ERF subfamily). The genes in class I all belong to ERF subfamily and contain a typical AP2 domain, the genes in Class II all belong to AP2 subfamily and contain two typical

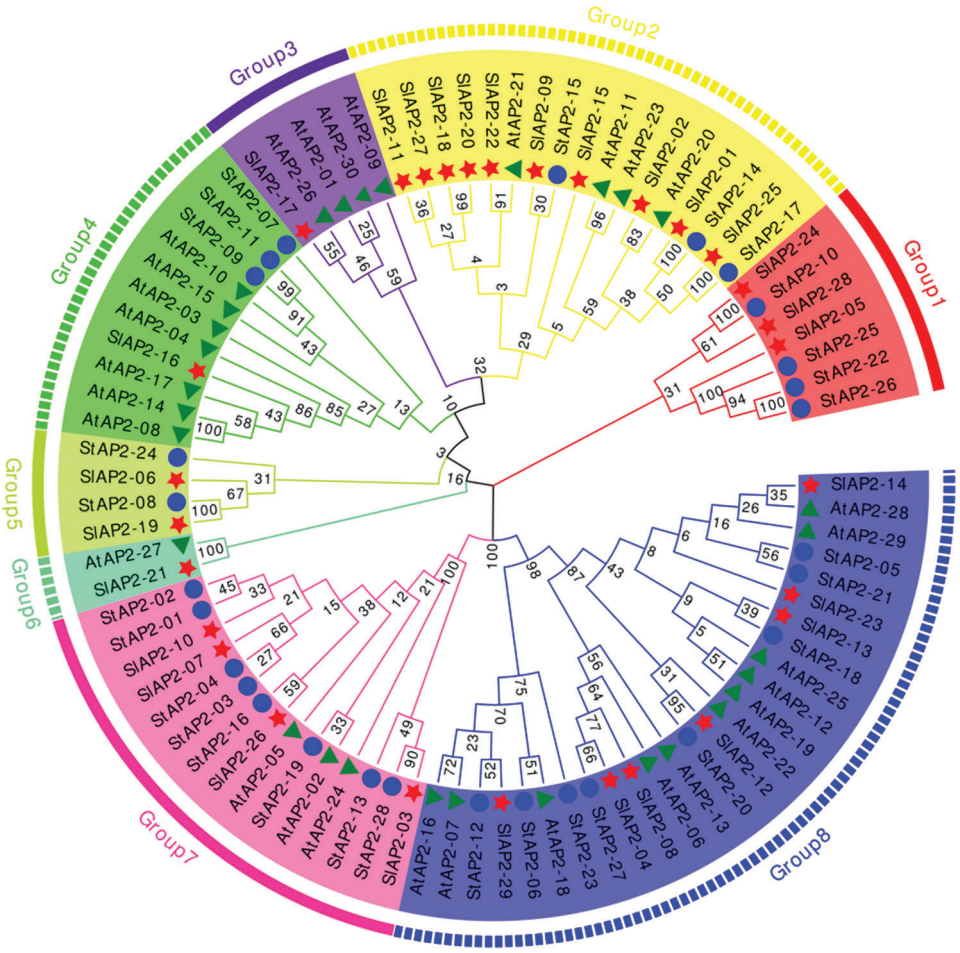

FIGURE 2. Phylogenetic analysis of AP2/ERF transcription factors from tomato and other plants.
The phylogenetic tree was constructed using 30, 28, and 29 AP2/ERF protein sequences from *Arabidopsis*, tomato, and potato, respectively. AP2/ERF members were divided into eight clades.

AP2 domains, and the genes in Class III contain an AP2 domain, but the 14 and 19 loci of the domain are both valine and L-aspartic acid, so they are both ERF subfamily and DERB subfamily.

The physical genome annotation files were obtained from genome sequences in phytozome database, and the 29 SlAP2/ERF genes identified were located on 10 chromosomes (Fig. 4). In general, SlAP2/ERF gene was distributed unevenly on the chromosomes. The highest chromosome 2 (6 genes) and the lowest chromosome 10 (1 gene), it suggests that the AP2/ERF gene family may have multiple functions in tomatoes. Meanwhile, it can be concluded that the genes clustered in the same subgroup in the phylogenetic tree are basically on the same chromosome.

Promoter analysis of AP2/ERF gene in tomato

In order to study homeopathic components of the promoter region of SlAP2/ERF genes, we intercepted the 2000 bp sequence upstream of transcription initiation for analysis by Plantcare online software. A total of 43 homeostatic components related to tissue-specific expression, plant hormone response, light response and stress response were screened (Fig. 5). We found that all SlAP2/ERF family genes contain multiple TATA box homeopathic response elements; secondly, there are many types of light-responsive elements, and they exist in 26 SlAP2/ERF gene families (except SlAP2-11, SlAP2-25 and SlAP2-28); in plant hormone response elements, CGTCA-motif and TGACG-motif are present in almost every gene; and we can also observe that some homeopathic elements are only present in a certain gene (SlAP2-15 in WUN-motif, TGA-box is in SlAP2-19, RY-element is in SlAP2-05, AT-rich element is in SlAP2-15, HD-Zip 1 is only in SlAP2-05). These results indicate that different SlAP2 genes regulate plant growth and development by responding to different homeopathic elements.

Analysis of homology model and construction of protein-protein interaction network

The Swiss-model online software was used to 3D model of the SlAP2/ERF protein family members (Fig. 6). The prediction model was based on the reported template, which was based on the maximum sequence fragment coverage, sequence identity, and credibility score of the test sequence. In addition, in order to clarify the similarity or difference of the generated models, the overlay structure is used to calculate the structural coverage. Structural coverage between the SlAP2 protein and the corresponding model sequence is more than 60%, indicating that the structural prediction of the SlAP2/ERF protein is relatively reliable. The three-dimensional model results show that these SlAP2/ERF proteins have similar tertiary structures, which mean that the SlAP2/ERF proteins may have evolved from the same ancestral sequence, or remained stable during long-term domestication after initial differentiation under the effect of purification selection.

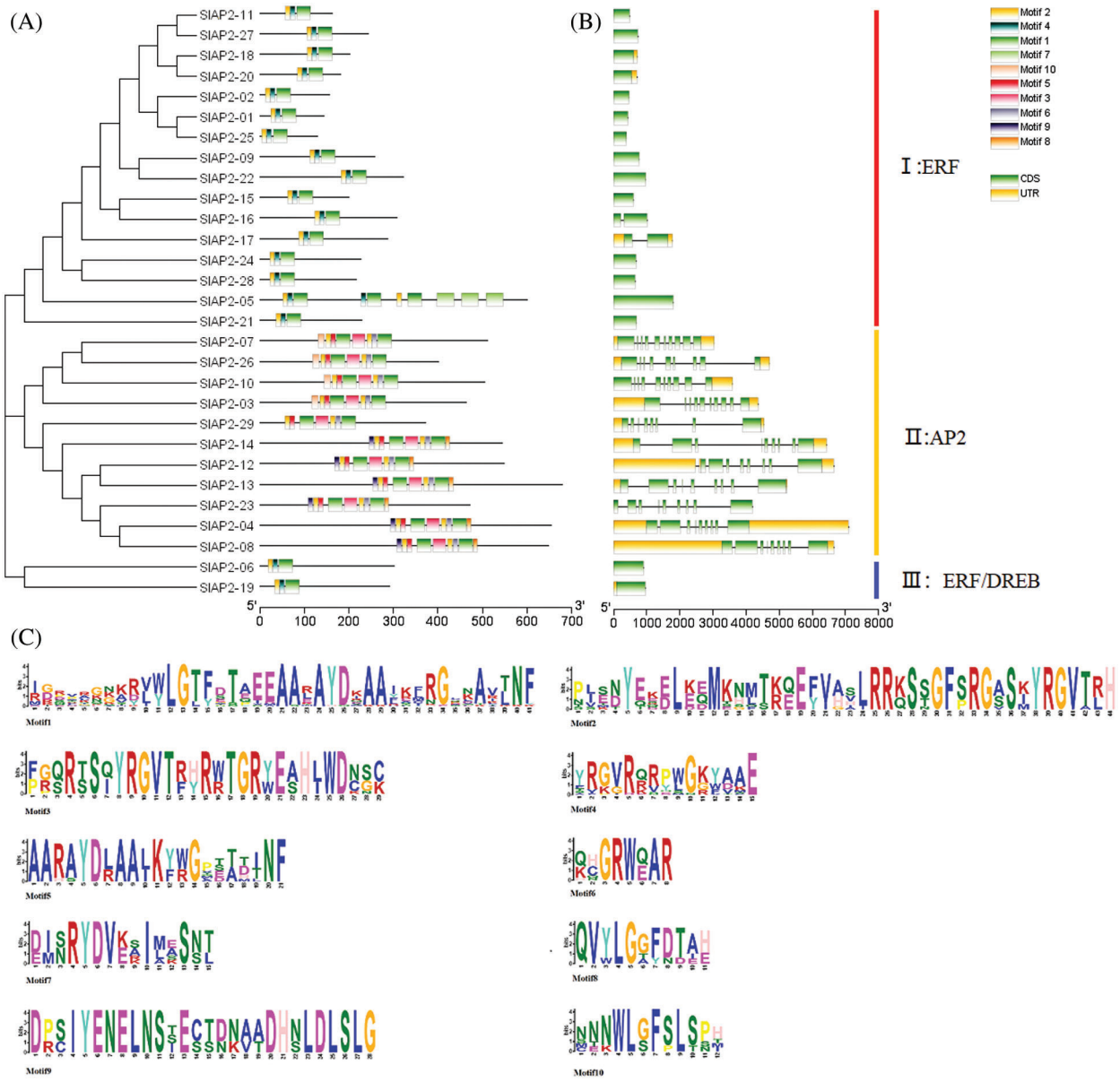

FIGURE 3. Gene structure and conserved protein domains of AP2/ERF.
A: The motif of AP2/ERF proteins. B: The intron-exon structures of SlAP2/ERF genes, including exons, introns. C: 10 conserved motifs sequence. Conserved motifs (1–10) are represented by different colored boxes while nonconserved sequences are shown by gray lines.

To further understand the function of SlAP2/ERF genes, we used String (http://STRING db.org/) to predict the protein-protein Interaction network of the SlAP2/ERF gene family in tomato and *Arabidopsis*. The results showed (Fig. 7) that all the SlAP2/ERF proteins appeared in the known interaction network of AP2 proteins in *Arabidopsis*, which indicated that there was a close relationship between *Arabidopsis* and tomato. The results also showed that the protein structure and sequence of *AtRAP2.11* were similar to those of two SlAP2/ERF proteins (SlAP2-06, SlAP2-19), and its function may be involved in regulation of gene expression through stress factors and components of the stress signal transduction pathway. *AtERF1* is similar to three SlAP2/ERF proteins (SlAP2-11, SlAP2-15, and SlAP2-27). It may act as a transcriptional activator, bind to the promoter elements related to GCC-box, participate in the regulation of gene expression during plant development, and be mediated by stress factors and components of stress signal transduction pathway. Similarly, *AtERF4* is structurally similar to SlAP2-24 and SlAP2-28, acting as a negative regulator of JA response to gene expression and fusarium oxysporum resistance to necrotizing fungal pathogens, and antagonizing TA inhibition of root elongation. *AtANT* is also involved in the initiation and development of organs, including floral organs, which maintain the cell's ability to be divided.

Gene ontology

SlAP2/ERF genes were annotated with Blast2GO. Among them, 17 genes were annotated in the GO database. According to the Fig. 8, SlAP2/ERF genes can be divided into three groups: biological process, molecular function, and cellular component. There were 17 genes (58.62%)

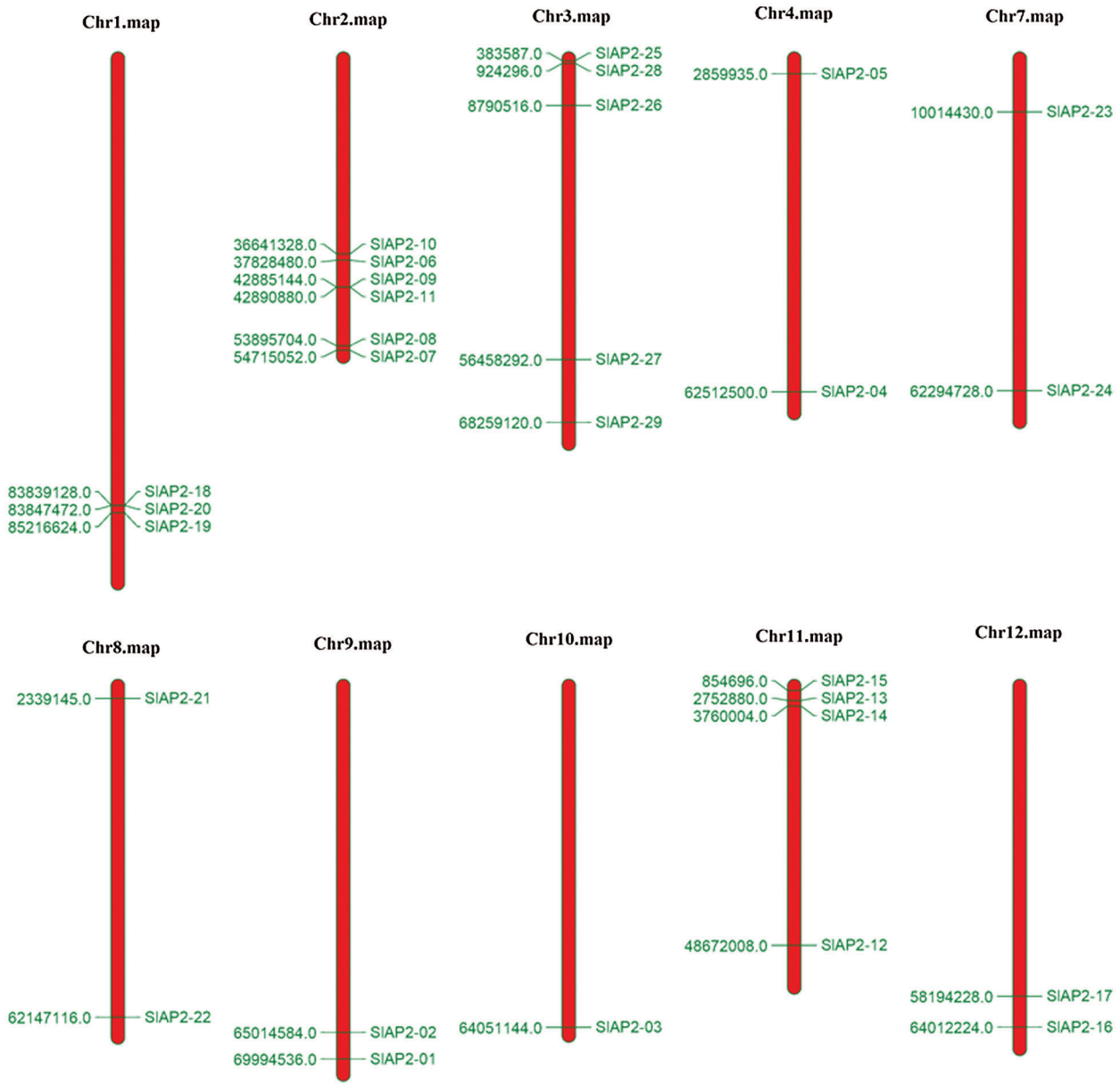

FIGURE 4. Distribution of SlAP2/ERF genes in chromosomes and the location of SlAP2/ERF transcription factors.

belonging to the biological process. Which transcription, DNA-templated (GO: 0006351) accounted for 16 genes (55.17%), followed by defense response (GO: 0006952) and ethylene-activated signaling pathway(GO: 0009873). In terms of molecular function, 13 SlAP2/ERF family genes are responsible for only two functions, DNA binding (GO: 0003677) and transcription factor activity (GO: 0003700). All 13 members of this family (44.82%) reside in the nucleus (GO: 0005634) and no other type sites exist in cellular component. We were able to observe an interesting phenomenon, where some SlAP2/ERF genes are involved in multiple biological processes, exist in the same cell component, and have multiple biological functions. These results suggest that the family genes are quite significant in nuclear development. For example, SlAP2-27 is located in the nucleus and has transcription factor activity and DNA binding, and it plays a key role in resistance to adversity, fruit ripening and activation of signaling pathways.

AP2/ERF gene expression profile analysis and expression analysis under biotic and abiotic stress by qRT-PCR

In order to analyze the expression profile of the AP2/ERF gene, we searched the tomato expression database for AP2/ERF gene expression abundance patterns in multiple tissues and organs (including flower, fruit on day 10 after flowering, fruit on day 20 after flowering, mature fruit, cotyledons, hypocotyls, meristems, mature leaves, roots, flower buds and young leaves). Heat map results (Fig. 9) showed that most of the genes have tissue expression specificity. For example, the SlAP2-03 gene is only highly expressed in the fruit on the 10th day after flowering, and SlAP2-26 is only specifically and highly expressed in mature fruit. It was observed that SlAP2-14 and SlAP2-16 were highly expressed in multiple tissues.

SlAP2/ERF transcription factors play a key role in abiotic stress, in order to further study the functions of these factors. We selected some specific genes through expression profile data in different tissues and homeopathic elements in the

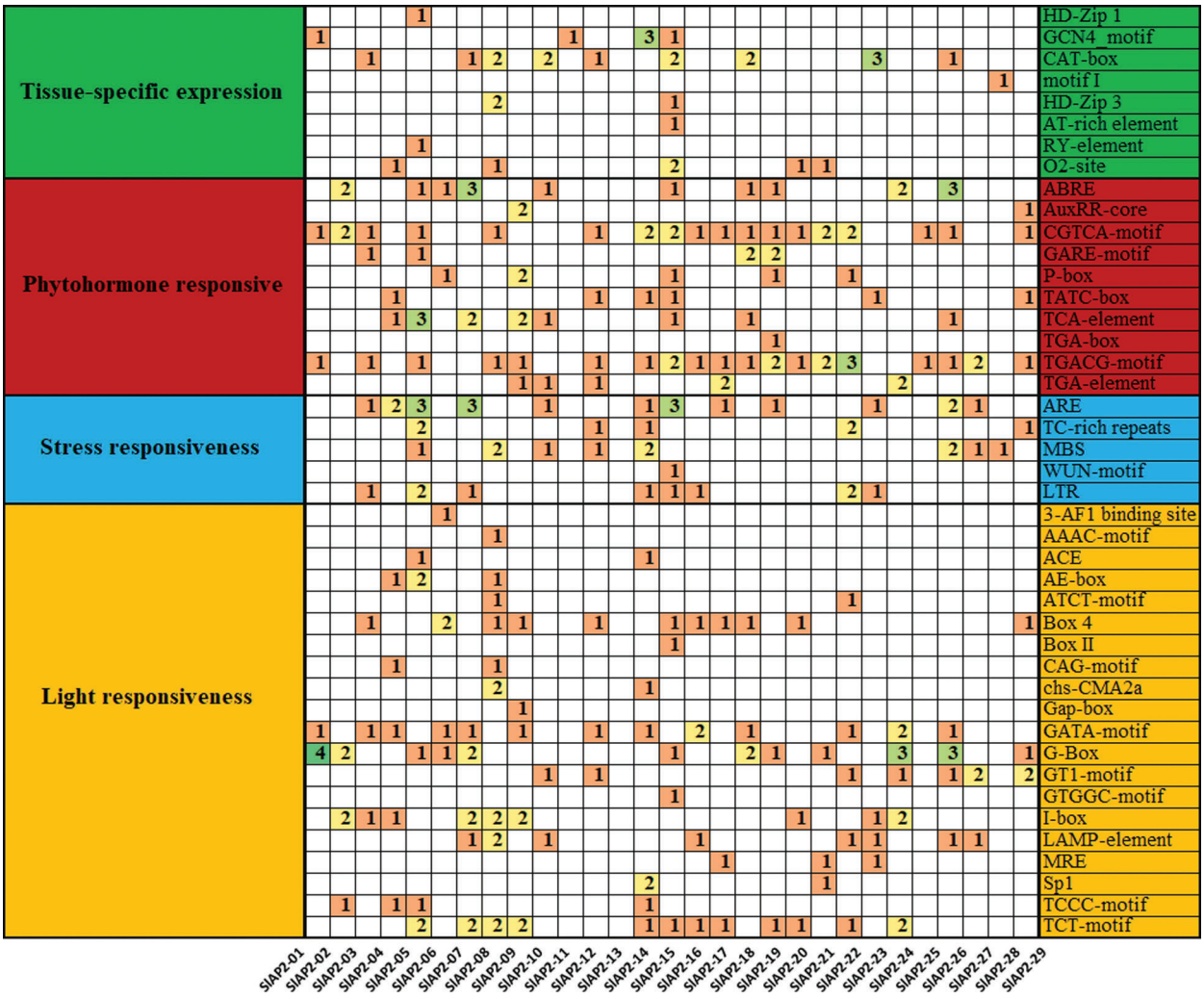

FIGURE 5. Analysis of cis-elements in promoters of 29 SlAP2/ERF genes.
The 2000 bp promoter regions (upstream DNA sequence of the 5′-UTR) of each SlAP2/ERF gene were analyzed for cis-elements. Different cis-elements were indicated by different colored diagrams.

promoter region. Under abiotic stress (20% PEG, 100 mmol/L Nacl, 100 umol/L ABA, 100 umol/L SA, 4°C low temperature), we studied the transcriptional abundance of these genes in tomato seedling leaves. The results showed that basically all 10 SlAP2/ERF genes responded to these abiotic stresses (Fig. 10). Among them, the expression of SlAP2-02 and SlAP2-22 changed significantly after Nacl treatment for 2 h and 4 h, and after low temperature treatment for 8 h and 12 h, while the expression levels were very low under other abiotic treatments. It is also interesting to observe that the expression pattern of most genes under five different treatments in the form of first increase and then decrease. When most of genes (SlAP2-05, SlAP2-07, SlAP2-08, SlAP2-09, SlAP2-14, SlAP2-15, SlAP2-24, and SlAP2-26) were treated for 8 h, the expression of the genes in the 5 treatments was significantly higher than that in the control. Meanwhile, under the Nacl treatment, we can clearly observe the expression pattern of some genes which increased first and then decreased (SlAP2-02, SlAP2-05, SlAP2-14, SlAP2-15, SlAP2-22, and SlAP2-26). After 8 h of low temperature treatment, the expression levels of 10 SlAP2/ERF genes were all high. Under the treatment of two hormones, the expression of most genes was very low (except that slap2-14, slap2-15 and slap2-26 were highly expressed at 12 h). Under PEG stress, the expression of some genes was significantly higher than that of the control group after 2 h. These results indicate that the SlAP2 genes could have a complex regulatory pattern in plants' resistance to abiotic stress.

Discussion

There are various kinds of environmental stresses in nature, such as biotic stress and abiotic stress; abiotic stress includes cold injury, salt injury, drought injury and flood injury. In order to resist these adverse abiotic environmental conditions, plants have developed various defense mechanisms at the molecular and physiological levels during their evolution (Cheong, 2003). There has been reported on abiotic stress in plants. On the molecular level, there are a large number of transcription factors in plants, which can respond to abiotic stresses and play a key role in regulating downstream targets. At the same time, it is quite important in plant growth, development, and response to various stresses (Mizoi *et al.*, 2012). With the gradual publication of the whole genome of some species, the identification and

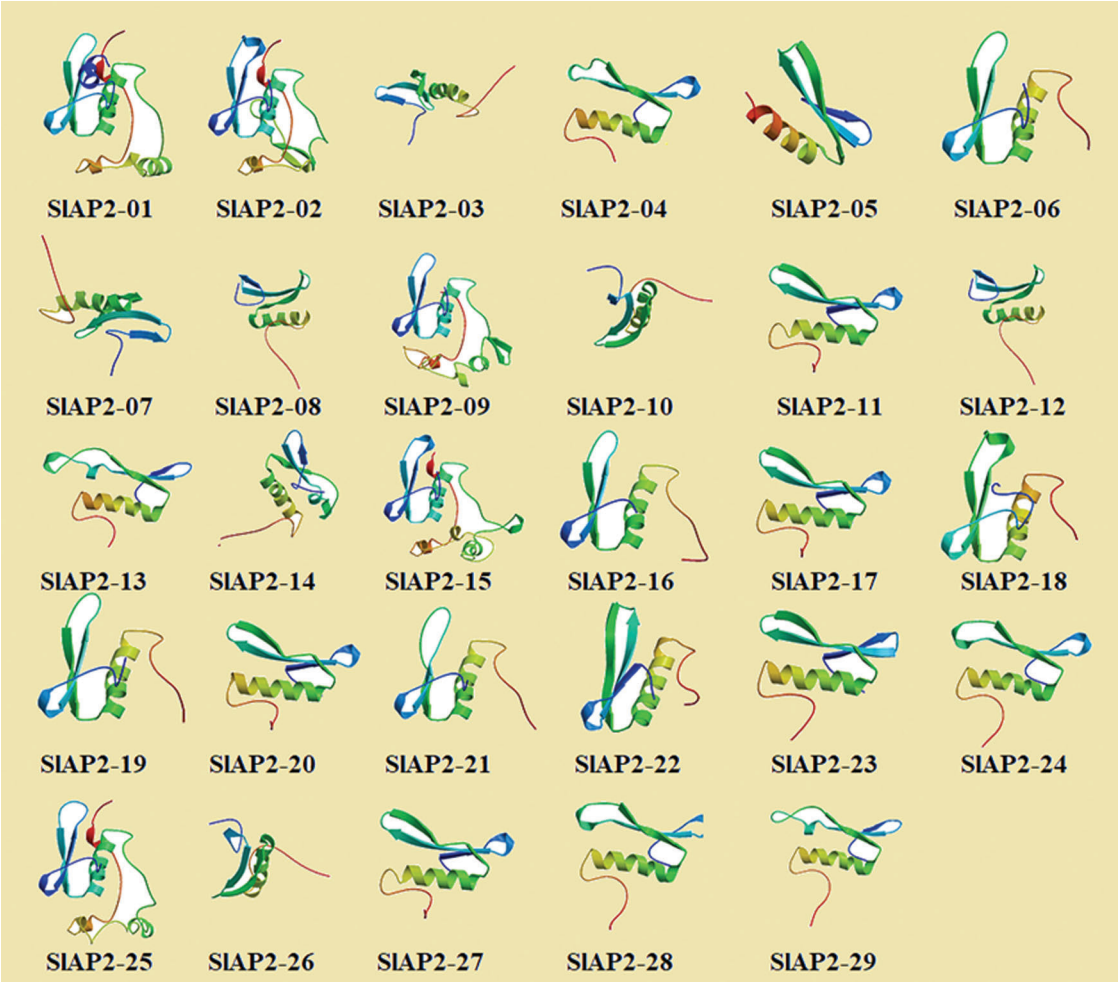

FIGURE 6. Predicted 3D models of tomato AP2/ERF proteins.
These models were constructed by the Swiss model server. Among them, template 2gcc.1.A were used in SlAP2-05; template 1gcc.1.c was used in SlAP2-23; and template 3gcc.1.A were used in SlAP2-17, SlAP2-20, SlAP2-22 and SlAP2-27; the rest of the members of SlAP2 were modelled with template 5wx91.A.

study of the classification, evolutionary characteristics, and functional prediction of gene families at the whole genome level have become the focus of genomics research. Genome-wide analysis of gene family is an effective method to describe the function of plant genes, which are helpful to study the evolution of genes and genomes. Many studies have been carried out on AP2/ERF transcription factors family and their functions in model plant *Arabidopsis thaliana*, and it has been found that AP2/ERF transcription factors family is mainly involved in plant growth and development, biotic and abiotic stress. While AP2/ERF transcription factors are widely distributed in plants, and because they can specifically bind to the cis-regulatory element in the promoter region to regulate the expression of downstream stress response genes, therefore, their role in plants under environmental stress is of great concern to researchers. AP2/ERF transcription factors have been studied in many plants, such as *Arabidopsis thaliana* (Guo *et al.*, 2005; Wang *et al.*, 2008), chickpea (Shukla *et al.*, 2006), rice (Rashid *et al.*, 2012), and maize (Lata *et al.*, 2014), and their classification and function have been reported.

However, the function of AP2/ERF in tomato has not been well studied. Therefore, we identified 29 typical AP2/ERF genes from tomato genome. AP2/ERF gene was analyzed from the aspects of basic physical and chemical properties, phylogeny, chromosome position, gene structure, protein sequence and three-level model, cis-regulatory elements, and expression level. We constructed phylogenetic trees from 87 AP2/ERF transcription factors identified in *Arabidopsis*, potato, and tomato. The results showed that 11 SlAP2/ERF, 16 StAP2/ERF, 14 AtAP2/ERF belonged to the typical AP2 subfamily, 15 SlAP2/ERF, 8 StAP2/ERF and 16 AtAP2/ERF belonged to the typical ERF subfamily, which indicated that the evolutionary mechanism of AP2/ERF transcription factors in tomato, potato and *Arabidopsis* was similar. Meanwhile, the number of AP2/ERF gene family mainly depends on the number of members of the ERF family (Shu *et al.*, 2016), which is consistent with our study.

The domains and motifs of transcription factors have an irreplaceable effect in many plant life activities, such as protein interactions, transcriptional activity, and DNA binding (Zhao *et al.*, 2019). In this study, we identified 10 conserved motifs in SlAP2/ERF. Among them, motif1 and motif2 are conserved in all identified sequences. The study of the motif may reveal the new biological function and regulation mechanism of tomato genes. It has been suggested that TFs, which shares

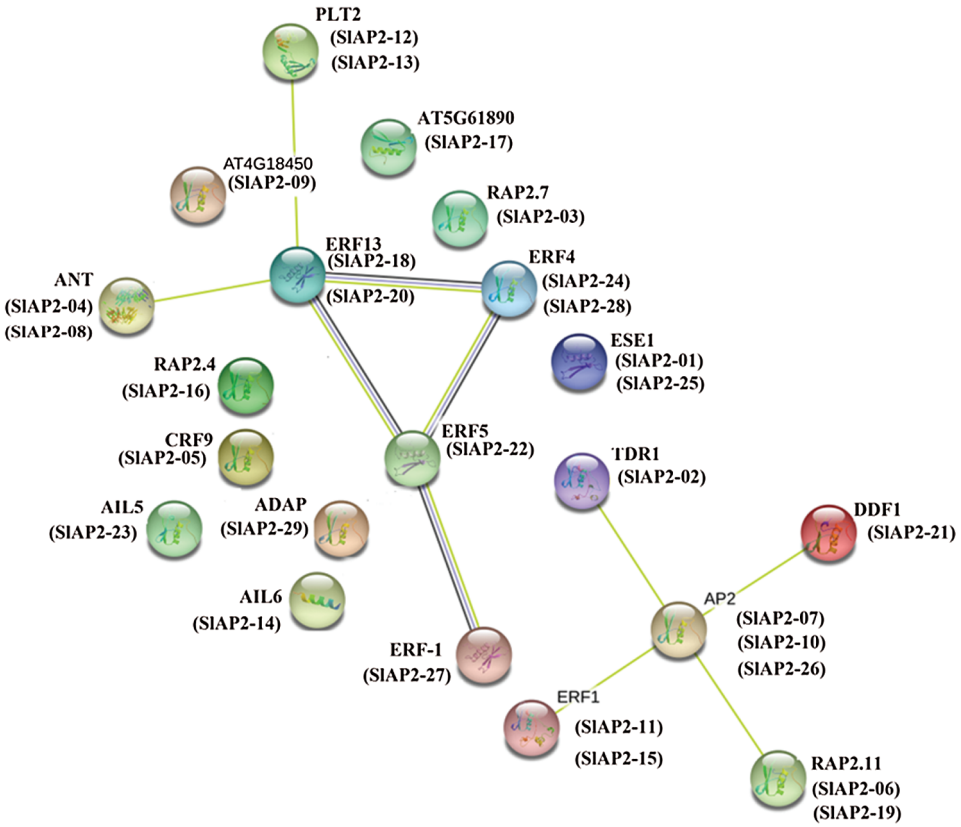

FIGURE 7. Prediction of SlAP2/ERF interaction network based on the homologous gene interaction in *Arabidopsis thaliana*.

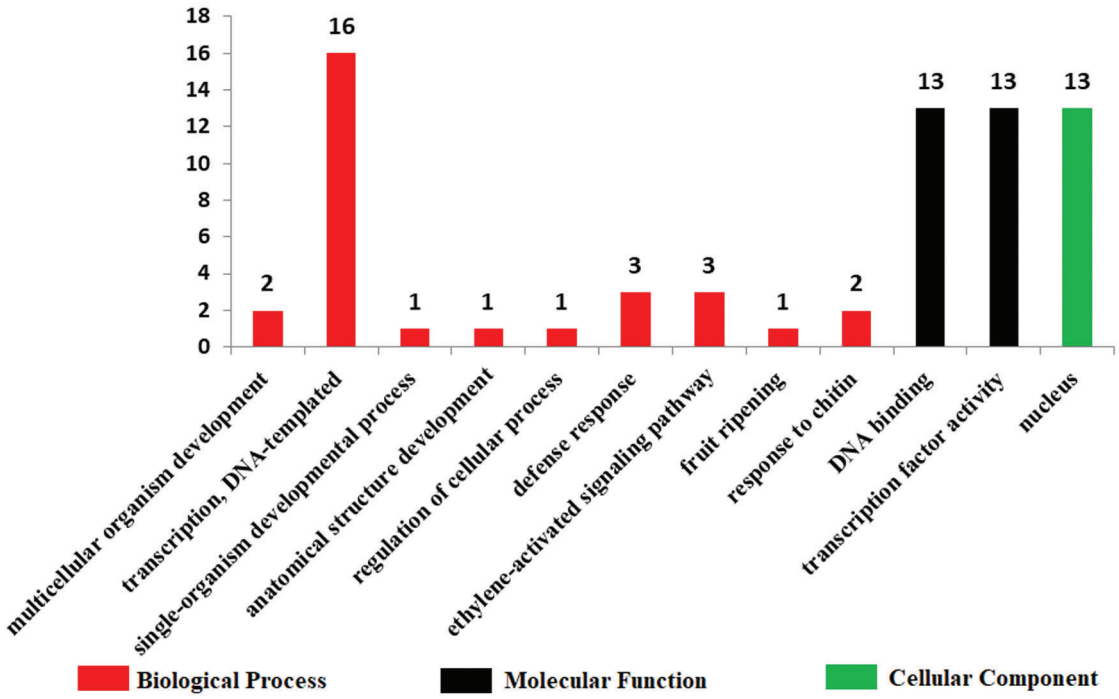

FIGURE 8. GO functional classifications of SlAP2/ERF.
The result of functional classification from three levels: biological process, cellular component, and molecular function.

unique motifs in a cluster, might have similar functions. For example, in the ERF protein family, the ERF related amphiphilic suppressor (EAR) motif is specifically present in the VIII genes, which is essential for the ERF protein's inhibitory function.

The study of gene structure can help us to further understand its function as well as the key information of evolution. In this study, we can conclude that genes of the same evolutionary branch have the same exon-intron structure. It is generally believed that genes with multiple

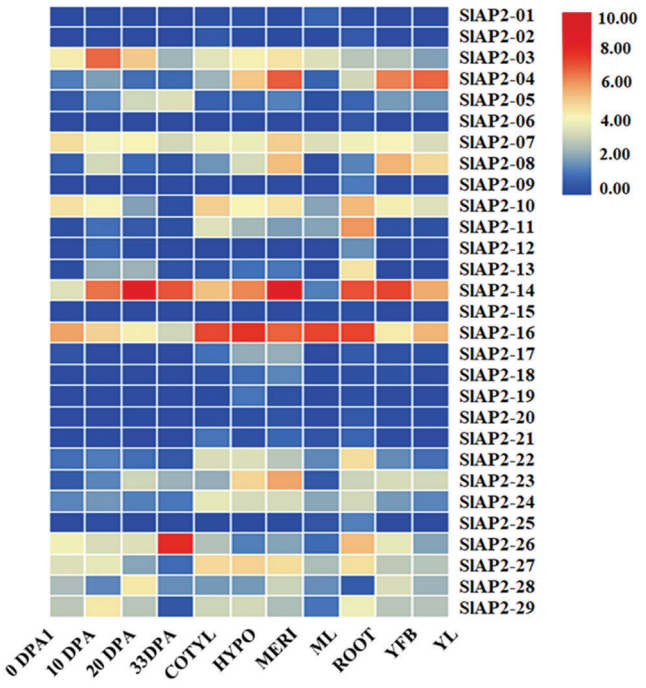

FIGURE 9. Relative expression analysis of SlAP2/ERF in different tissues.
Including flower, fruit on day 10 after flowering, fruit on day 20 after flowering, mature fruit, cotyledons, hypocotyls, meristems, mature leaves, roots, flower buds and young leaves. The $2^{-\Delta\Delta CT}$ method was used to be an evaluation of the relative expression. The heat map was drawn in row scale-transformed expression values. Red or green colors represent the difference in expression levels in each sample, respectively.

introns are conservative, while genes without introns are less conservative (Kaufmann *et al.*, 2005). We found that 15 ERF/DREB subfamily genes and one gene (SlAP2-05) containing three AP2 domains had no intron structure, and 2 ERF/DREB subfamily members had only one intron domain, and only 3 types of conserved sequences were found in them. The number of introns in the typical 11 AP2 subfamily members is between 7 and 9. Previous studies have shown that introns gain and lose as a result of stress selection during the long-term evolution of plants, and genes gradually form various types of intron-exon structures and perform different functions (Wang *et al.*, 2016). The results showed that the SlAP2/ERF family evolved into different subfamilies to regulate various life activities during the long-term development of the tomato genome.

Promoter, as the key DNA sequence that regulates gene expression, plays an important role in the initiation of transcription, and the number of promoters directly affects the efficiency of downstream gene transcription. The promoter region of the cis-regulatory element is irreplaceable in regulating gene expression and plant resistance to adverse environmental conditions, and the TATA-box is the key cis-regulatory element, it is essential for the correct expression of most eukaryotic genes and determines the transcription initiation site and its initiation frequency. In the present study, we identified several homeostatic components in the upstream promoter region of the SlAP2/ERF genes that respond to plant hormones, stress and tissue-specific expression (Fig. 4). This suggests that the function of SlAP2/ERF genes is mediated by a complex regulatory network. Previous study showed that AP2/ERF protein could bind to GCC-box motif through ERF domain and regulated the expression of target gene under stress (Shigyo *et al.*, 2006; Lata and Prasad, 2011). In this study, the homology of SlAP2-01 in *Arabidopsis thaliana*, *ESE1*, is linked to the promoter of GCC, which is involved in the regulation of gene expression. *ESE1*, as the homologous gene of SlAP2-01 in *Arabidopsis*, it is involved in the regulation of gene expression by connecting to the promoter element of GCC frame, which is related to the pathogenesis of *Arabidopsis*, and studies showed that *AtESE1* could enhance the salt tolerance (Zhang *et al.*, 2011). So SlAP2-01 may have *ESE1*-like features. qRT-PCR results and SlAP2/ERF gene expression profiles in different tissues showed that SlAP2/ERF gene expression changed significantly under different abiotic stresses and hormone treatments. Tissue expression analysis showed that AP2/ERF genes were expressed in tomato tissues, and the expression of AP2/ERF genes was different in different tissues. Most genes are highly expressed in fruit and flowers, which is consistent with previous reports of ERF proteins involved in fruit development (Licausi *et al.*, 2013), which may indicate that ERF genes are also involved in the development of flower organs and fruit. Abiotic stress analysis of 10 genes found that most genes were induced by stress. After 4 h of low temperature treatment, the expression of 10 genes was significantly higher than that of the control group, which is consistent with the results of previous studies on ERF participation in cold stress (Gao *et al.*, 2014). Seven genes were significantly upregulated at 4th of salt treatment; most genes were induced by ABA and SA treatment. This provides an important basis for further studying the function and mechanism of tomato ERF transcription factor family genes. In this study, we studied potential abiotic stress response elements within 2000 bp upstream of the start codon ATG, including ABA, low temperature, drought stress, and salt stress regulatory elements. The results show that all SlAP2 genes contain these stress response elements. Based on the expression profile of qRT-PCR, we can observe SlAP2-05 and SlAP2-22 genes with two low temperature response elements (LTR), SlAP2-02, SlAP2-07, and SlAP2-26 genes containing multiple ABA response elements (ABRE), SlAP2-08, SlAP2-14 and SlAP2-26 genes with two drought response elements (MBS), SlAP2-05, SlAP2-07 and SlAP2-09 genes with several salicylic acid response elements (TCA-element), their expression levels were significantly up- or down-regulated, which was consistent with the analysis of regulatory elements and also with the results of HD-I ZIP Gene Subfamily in *Nicotiana tabacum* (Li *et al.*, 2019). Meanwhile, these results indicated that gene expression is a complex biological process, which is regulated by many factors. Therefore, the expression mechanism of these tomato SlAP2/ERF genes needs further study.

The study of tertiary structure and protein-protein interaction can help us to understand the function of SlAP2/ERF members more clearly. The result of tertiary modeling shows that genes located in the same branch have similar tertiary structure (Fig. 5). Although the AP2/ERF gene

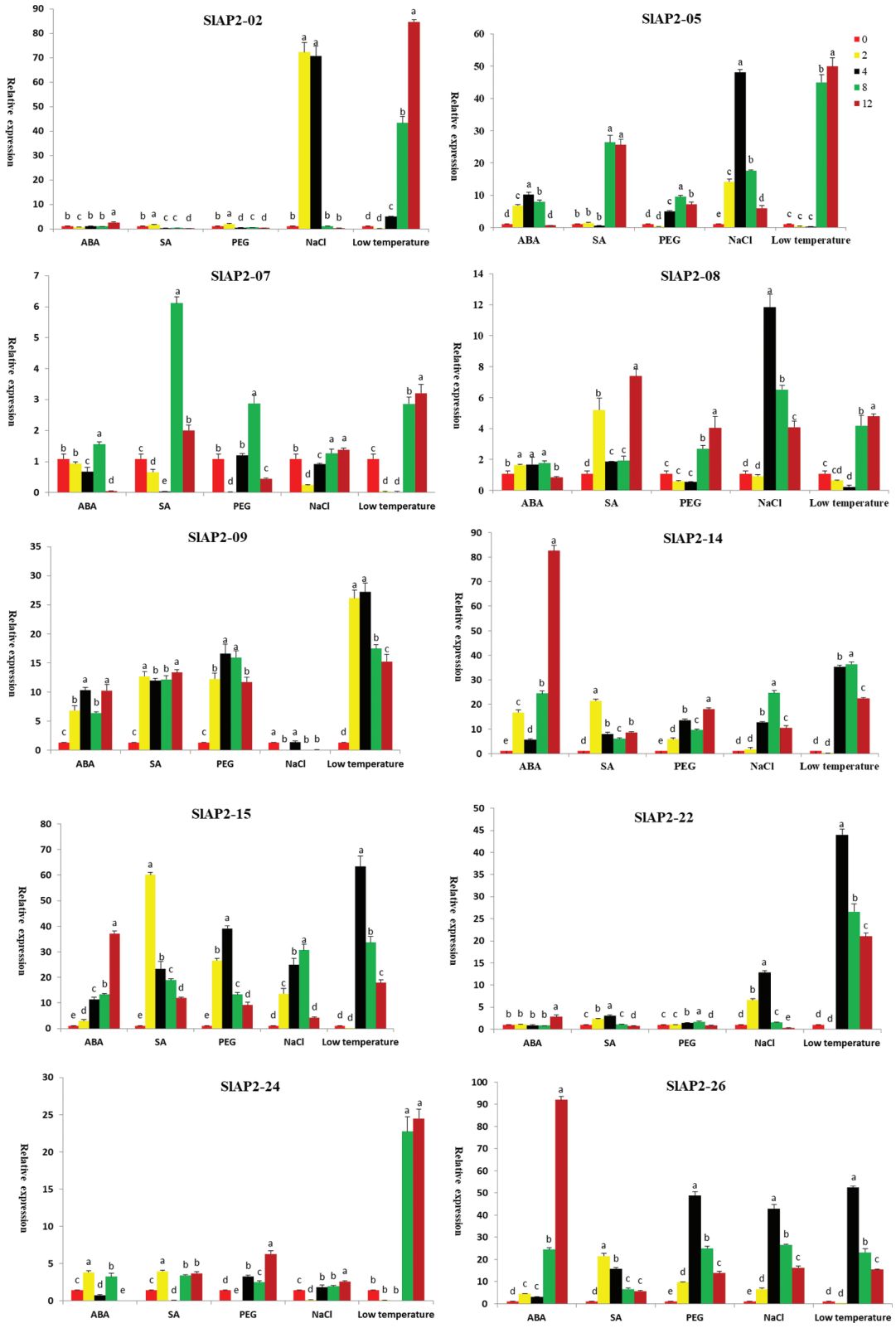

FIGURE 10. Expression patterns of 10 SlAP2/ERF under different treatments (20% PEG, 100 mmol/L Nacl, 100 umol/L ABA, 100 umol/L SA, 4°C low temperature).

Bars represent the mean values of three replicates ± standard deviation (SD). Red represents control treatment, yellow represents 2 h after processing, black represents 4 h after processing, green represents 8 h after processing, and red brown represents 12 h after processing.

family in *Arabidopsis* has been shown to influence many biological processes, relatively little has been studied in tomato, so we constructed a network of AP2/ERF protein interactions between tomato and *Arabidopsis*. Previous results showed that *AtERF4* negatively regulated iron deficiency response in *Arabidopsis* (Liu *et al.*, 2017), and *AtERF13* endow *Arabidopsis* with ABA hypersensitivity (Lee *et al.*, 2010). Therefore, the corresponding proteins of

tomato, SlAP2-24, SlAP2-28 and AtERF4, and SlAP2-18, SlAP2-20 and *AtERF13*, have similar functions. The interaction network showed that SlAP2/ERF genes play an irreplaceable role in the development of tomato and response to abiotic and biotic stresses. Go analysis showed that SlAP2/ERF genes were widely involved in many biological processes, and were distributed in the nucleus of cells, and had DNA binding and transcription factor activity. The results showed that *CrERF5* positively regulated the biosynthesis of bisindole alkaloids and their precursors in *Catharanthus roseus* (Pan *et al*., 2019), and the AP2 / ERF family transcription factor GSERF71 from soybean was a DNA binding protein, which positively regulated the tolerance alkaline stress tolerance in *Arabidopsis* (Zhang *et al*., 2018).

Conclusion

In conclusion, we analyzed the AP2/ERF gene family of tomato at the whole genome level, and identified 29 SlAP2/ERF genes, and divided them into subfamilies. At the same time, its basic physical and chemical properties were analyzed, and the phylogenetic tree was constructed with 30 AP2/ERF proteins from *Arabidopsis thaliana* and 28 AP2/ERF proteins from potato to ensure their conservative homology. We also analyzed the chromosomal location, conserved motifs, gene structure, promoter elements, protein tertiary structure and protein interaction network of SlAP2/ERF transcription factor. Meanwhile, the expression level of SlAP2/ERF in different tissues and under different stress was detected by transcriptome and q-RT-PCR. These results provided comprehensive information for further study of the function of the SlAP2/ERF gene family.

Acknowledgement: We would like to thank the reviewers and editors for careful reading and helpful comments on this manuscript.

Availability of Data and Materials: All data generated or analysed during this study are included in this published article.

Funding Statement: This project was funded by the National Natural Science Foundation of China (31560663, 32060401).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

Aya K, Hobo T, Sato-Izawa K, Ueguchi-Tanaka M, Kitano H, Matsuoka M (2014). A novel AP2-type transcription factor, SMALL ORGAN SIZE1, controls organ size downstream of an auxin signaling pathway. *Plant and Cell Physiology* **55**: 897–912. DOI 10.1093/pcp/pcu023.

Bailey TL, Williams N, Misleh C, Li WW (2006). MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Research* **34**: W369–W373. DOI 10.1093/nar/gkl198.

Boutilier K, Offringa R, Sharma VK, Kieft H, Ouellet T, Zhang L, Hattori J, Liu CM, Lammeren A, Miki B, Custer J, Lookeren Campagen M (2002). Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *Plant Cell* **14**: 1737–1749. DOI 10.1105/tpc.001941.

Cheong YH (2003). BWMK1, a rice mitogen-activated protein kinase, locates in the nucleus and mediates pathogenesis-related gene expression by activation of a transcription factor. *Plant Physiology* **132**: 1961–1972.

Chinnusamy V, Schumaker K, Zhu JK (2003). Molecular genetic perspectives on cross-talk and specificity in abiotic stress signalling in plants. *Journal of Experimental Botany* **55**: 225–236. DOI 10.1093/jxb/erh005.

Consortium TTG (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **485**: 635–641. DOI 10.1038/nature11119.

Del Campo EM, Casano LM, Barreno E (2013). Evolutionary implications of intron-exon distribution and the properties and sequences of the RPL10A gene in eukaryotes. *Molecular Phylogenetics and Evolution* **66**: 857–867. DOI 10.1016/j.ympev.2012.11.013.

Finn RD, Tate J, Mistry J (2007). The Pfam protein families database. *Nucleic Acids Research* **36**: D281–D288. DOI 10.1093/nar/gkm960.

Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C, Jensen LJ (2012). STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Research* **41**: D808–D815. DOI 10.1093/nar/gks1094.

Gao JS, Hu L, Xie P, Meng Y, Cai YP, Lin Y (2014). Isolation and molecular characterization of an ethylene response factor NtERF1-1 in Nicotiana tabacum cv. Xanthi. *Journal of Biosciences* **39**: 887–897. DOI 10.1007/s12038-014-9473-5.

Gilmour SJ, Sebolt AM, Salazar MP, Everard JD, Thomashow MF (2000). Overexpression of the *Arabidopsis* CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation. *Plant Physiology* **124**: 1854–1865. DOI 10.1104/pp.124.4.1854.

Guo A, He K, Liu D, Bai SN, Gu XC, Wei LP, Luo JC (2005). DATF: a database of *Arabidopsis* transcription factors. *Bioinformatics* **21**: 2568–2569. DOI 10.1093/bioinformatics/bti334.

Gutterson N, Reuber TL (2004). Regulation of disease resistance pathway by AP2/ERF transcription factors. *Current Opinion in Plant Biology* **7**: 26857–26861.

Hobert O (2008). Gene regulation by transcription factors and micro-RNAs. *Science* **319**: 1785–1786. DOI 10.1126/science.1151651.

Hu B, Jin JP, Guo AY, Zhang H, Luo JC, Gao G (2015). GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics* **31**: 1296–1297. DOI 10.1093/bioinformatics/btu817.

Hu YX, Wang YX, Liu XF, Li JY (2004). *Arabidopsis* RAV1 is down-regulated by brassinosteroid and may act as a negative regulator during plant development. *Cell Research* **14**: 8–15. DOI 10.1038/sj.cr.7290197.

Jiang N, Cui J, Shi Y, Yang G, Zhou X (2019). Tomato lncRNA23468 functions as a competing endogenous RNA to modulate NBS-LRR genes by decoying miR482b in 369 tomato-*Phytophthora infestans* interaction. *Horticulture Research* **6**: 215. DOI 10.1038/s41438-018-0096-0.

Jofuku KD, Boer BG, Van MM, Okamuro JK (1994). Control of *Arabidopsis* flower and seed development by the homeotic gene APETALA2. *Plant Cell* **6**: 1211–1225.

Kaufmann K, Melzer R, Theissen G (2005). MIKC-type MADS-domain proteins: Structural modularity, protein interactions

and network evolution in land plants. *Gene* **347**: 183–198. DOI 10.1016/j.gene.2004.12.014.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10**: 845–858. DOI 10.1038/nprot.2015.053.

Kudo K, Oi T, Uno Y (2014). Functional characterization and expression profiling of a *DREB2*-type gene from lettuce (*Lactuca sativa* L.). *Plant Cell, Tissue and Organ Culture* **116**: 97–109. DOI 10.1007/s11240-013-0386-z.

Kunst L, Klenz JE, Martinezzapater J, Haughn GW (1989). AP2 gene determines the identity of Perianth organs in flowers of *Arabidopsis* thaliana. *Plant Cell* **1**: 1195–1208. DOI 10.2307/3868917.

Lata C, Mishra AK, Muthamilarasan M, Bonthala VS, Khan Y, Manoj P (2014). Genome-wide investigation and expression profiling of AP2/ERF transcription factor superfamily in foxtail millet (*Setaria italica L.*). *PLoS One* **9**: e113092. DOI 10.1371/journal.pone.0113092.

Lata C, Prasad M (2011). Role of DREBs in regulation of abiotic stress responses in plants. *Journal of Experimental Botany* **62**: 4731–4748. DOI 10.1093/jxb/err210.

Lee SJ, Park JH, Lee MH, Yu JH, Kim SY (2010). Isolation and functional characterization of CE1 binding proteins. *BMC Plant Biology* **10**: 277. DOI 10.1186/1471-2229-10-277.

Lescot M, Dehais P, Thijs G (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Research* **30**: 325–327. DOI 10.1093/nar/30.1.325.

Letunic I, Doerks T, Bork P (2015). SMART: Recent updates, new developments and status in 2015. *Nucleic Acids Research* **43**: D257–D260. DOI 10.1093/nar/gku949.

Li CW, Su RC, Cheng CP, Sanjaya-You SJ, Hsieh TH, Chao TC, Chan MT (2011). Tomato RAV transcription factor is a pivotal modulator involved in the AP2/EREBP-mediated defense pathway. *Plant Physiology* **156**: 213–227. DOI 10.1104/pp.111.174268.

Li H, Wang Y, Wu M, Li L, Li C, Han ZP, Yuan JY, Chen CB, Song WQ, Wang CG (2017). Genome-wide identification of AP2/ERF transcription factors in cauliflower and expression profiling of the ERF family under salt and drought stresses. *Frontiers in Plant Science* **8**: 1051. DOI 10.3389/fpls.2017.00946.

Li XL, Tao S, Wei SW, Ming ML, Huang XS, Zhang SL, Wu J (2018). The mining and evolutionary investigation of AP2/ERF genes in pear (Pyrus). *BMC Plant Biology* **18**: 2105. DOI 10.1186/s12870-018-1265-x.

Li YY, Bai BH, Wen F, Zhao M, Xia QY, Yang DH, Wang GH (2019). Genome-wide identification and expression analysis of HD-ZIPI gene subfamily in *Nicotiana tabacum*. *Genes (Basel)* **10**: 575.

Licausi F, Ohme-Takagi M, Perata P (2013). APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors: mediators of stress responses and developmental programs. *New Phytologist* **199**: 639–649. DOI 10.1111/nph.12291.

Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1998). Two transcription factors, ERF1 and ERF2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought and low temperature responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell* **10**: 1391–1406. DOI 10.1105/tpc.10.8.1391.

Liu W, Karemera NJU, Wu T, Yang YF, Zhang XZ, Xu XF, Wang Y, Han ZH (2017). The ethylene response factor AtERF4 negatively regulates the iron deficiency response in *Arabidopsis* thaliana. *PLoS One* **12**: e0186580. DOI 10.1371/journal.pone.0186580.

Liu ZN, Kong LJ, Zhang M, Lv YX, Liu YP, Zou MH, Lu G, Cao JS, Yu XL (2013). Genome-wide identification, phylogeny, evolution and expression patterns of AP2/ERF genes and cytokinin response factors in *Brassica rapa* ssp. pekinensis. *PLoS One* **8**: e83444. DOI 10.1371/journal.pone.0083444.

Livak KJ, Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods* **25**: 402–408. DOI 10.1006/meth.2001.1262.

Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K (2012). AP2/ERF family transcription factors in plant abiotic stress responses. *Biochimica et Biophysica Acta (BBA)—Gene Regulatory Mechanisms* **1819**: 86–96. DOI 10.1016/j.bbagrm.2011.08.004.

Nakano T, Suzuki K, Fujimura T, Shinshi H (2006). Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. *Plant Physiology* **140**: 411–432. DOI 10.1104/pp.105.073783.

Novillo F, Medina J, Rodríguez-Franco M, Neuhaus G, Salinas J (2011). Genetic analysis reveals a complex regulatory network modulating CBF gene expression and *Arabidopsis* response to abiotic stress. *Journal of Experimental Botany* **63**: 293–304. DOI 10.1093/jxb/err279.

Pan IC, Li CW, Su RC, Cheng CP, Lin CS, Chan MT (2010). Ectopic expression of an EAR motif deletion mutant of SlERF3 enhances tolerance to salt stress and *Ralstonia solanacearum* in tomato. *Planta* **232**: 1075–1086. DOI 10.1007/s00425-010-1235-5.

Pan QF, Wang CY, Xiong ZW, Wang H, Fu XQ, Shen Q, Peng BW, Ma YN, Sun XF, Tang KX (2019). CrERF5, an AP2/ERF transcription factor, positively regulates the biosynthesis of bisindole alkaloids and their precursors in *Catharanthus roseus*. *Frontiers in Plant Science* **10**: 531. DOI 10.3389/fpls.2019.00931.

Qin F, Sakuma Y, Li J, Liu Q, Li YQ, Shinozaki K, Shinozaki KY (2004). Cloning and functional analysis of a novel ERF1/CBF transcription factor involved in cold-responsive gene expression in *Zea mays* L. *Plant and Cell Physiology* **45**: 1042–1052. DOI 10.1093/pcp/pch118.

Quan R, Hu S, Zhang Z, Zhang H, Huang R (2010). Overexpression of an ERF transcription factor *TSRF1* improves rice drought tolerance. *Plant Biotechnology Journal* **8**: 476–488. DOI 10.1111/j.1467-7652.2009.00492.x.

Rashid M, Guang YH, Guang XY, Hussain J, Xu Y (2012). AP2/ERF transcription factor in rice: genome-wide canvas and syntenic relationships between monocots and eudicots. *Evolutionary Bioinformatics* **8**: EBO.S9369. DOI 10.4137/EBO.S9369.

Rogozin IB, Sverdlov AV, Babenko VN, Koonin EV (2005). Analysis of evolution of exon-intron structure of eukaryotic genes. *Briefings in Bioinformatics* **6**: 118–134. DOI 10.1093/bib/6.2.118.

Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K (2002). DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* ERFs, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochemical and Biophysical Research Communications* **290**: 998–1009. DOI 10.1006/bbrc.2001.6299.

Sakuma Y, Maruyama K, Osakabe Y, Qin F, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2006). Functional analysis of an *Arabidopsis* transcription factor, ERF2A, involved in

drought-responsive gene expression. *Plant Cell* **18**: 1292–1309. DOI 10.1105/tpc.105.035881.

Shigyo M, Hasebe M, Ito M (2006). Molecular evolution of the AP2 subfamily. *Gene* **366**: 256–265. DOI 10.1016/j.gene.2005.08.009.

Shu Y, Liu Y, Zhang J, Song L, Guo C (2016). Genome-wide analysis of the AP2/ERF superfamily genes and their responses to abiotic stress in *Medicago truncatula*. *Frontiers in Plant Science* **6**: 1051. DOI 10.3389/fpls.2015.01247.

Shukla RK, Raha S, Tripathi V, Chattopadhyay D (2006). Expression of CAP2, an APETALA2-family transcription factor from chickpea, enhances growth and tolerance to dehydration and salt stress in transgenic tobacco. *Plant Physiology* **142**: 113–123. DOI 10.1104/pp.106.081752.

Singh K, Foley RC, Onate-Sanchez L (2002). Transcription factors in plant defense and stress responses. *Current Opinion in Plant Biology* **5**: 430–436. DOI 10.1016/S1369-5266(02)00289-3.

Sohn KH, Lee SC, Jung HW, Hong JK, Kook Hwang BK (2006). Expression and functional roles of the pepper pathogen-induced transcription factor RAV1 in bacterial disease resistance, and drought and salt stress tolerance. *Plant Molecular Biology* **61**: 897–915. DOI 10.1007/s11103-006-0057-0.

Soltész A, Smedley M, Vashegyi I, Galiba G, Harwood W, Vagufalvi A (2013). Transgenic barley lines prove the involvement of TaCBF14 and TaCBF15 in the cold acclimation process and in frost tolerance. *Journal of Experimental Botany* **64**: 1849–1862. DOI 10.1093/jxb/ert050.

Song X, Ying L, Hou X (2013). Genome-wide analysis of the AP2/ERF transcription factor superfamily in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *BMC Genomics* **14**: 573. DOI 10.1186/1471-2164-14-573.

Song XM, Huang ZN, Duan WK (2014). Genome-wide analysis of the bHLH transcription factor family in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *Molecular Genetics and Genomics* **289**: 77–91. DOI 10.1007/s00438-013-0791-3.

Thamilarasan SK, Park JI, Jung H, Nou IS (2014). Genome-wide analysis of the distribution of AP2/ERF transcription factors reveals duplication and CBFs genes elucidate their potential function in Brassica oleracea. *BMC Genomics* **15**: 422. DOI 10.1186/1471-2164-15-422.

Venkatesh J, Yu J, Park SW (2013). Genome-wide analysis and expression profiling of the *Solanum tuberosum* aquaporins. *Plant Physiology and Biochemistry* **73**: 392–404. DOI 10.1016/j.plaphy.2013.10.025.

Voorrips RE (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *Journal of Heredity* **93**: 77–78. DOI 10.1093/jhered/93.1.77.

Wang M, Yue H, Feng KW, Deng PC, Song WN, Nie XJ (2016). Genome-wide identification, phylogeny and expressional profiles of mitogen activated protein kinase kinase kinase (MAPKKK) gene family in bread wheat (*Triticum aestivum* L.). *BMC Genomics* **17**: 749. DOI 10.1186/s12864-016-2993-7.

Wang QY, Guan YC, Wu YR, Chen HL, Chen F, Chu CC (2008). Overexpression of a rice OsERF1F gene increases salt, drought, and low temperature tolerance in both *Arabidopsis* and rice. *Plant Molecular Biology* **67**: 589–602. DOI 10.1007/s11103-008-9340-6.

Wu LJ, Chen XL, Ren HY, Zhang ZJ, Zhang HW, Wang JY, Wang XC, Huang RF (2007). ERF protein JERF1 that transcriptionally modulates the expression of abscisic acid biosynthesis-related gene enhances the tolerance under salinity and cold in tobacco. *Planta* **226**: 815–825. DOI 10.1007/s00425-007-0528-9.

Wu ZJ, Li XH, Liu ZW, Li H, Wang YX, Zhuang J (2015). Transcriptome-based discovery of AP2/ERF transcription factors related to temperature stress in tea plant (*Camellia sinensis*). *Functional & Integrative Genomics* **15**: 741–752. DOI 10.1007/s10142-015-0457-9.

Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, Bolund L, Wang J (2006). WEGO: a web tool for plotting GO annotations. *Nucleic Acids Research* **34**: W293–W297. DOI 10.1093/nar/gkl031.

Zhang B, Su L, Hu B, Li L (2018). Expression of AhDREB1, an AP2/ERF transcription factor gene from peanut, is affected by histone acetylation and increases abscisic acid sensitivity and tolerance to osmotic stress in arabidopsis. *International Journal of Molecular Sciences* **19**: 1441. DOI 10.3390/ijms19051441.

Zhang H, Liu W, Wan L, Li F, Dai L, Li D, Zhang Z, Huang R (2010). Functional analyses of ethylene response factor JERF3 with the aim of improving tolerance to drought and osmotic stress in transgenic rice. *Transgenic Research* **19**: 809–818. DOI 10.1007/s11248-009-9357-x.

Zhang LX, Li ZF, Quan RD, Li GJ, Wang RG, Huang RF (2011). An AP2 domain-containing gene, ESE1, targeted by the ethylene signaling component EIN3 is important for the salt response in *Arabidopsis*. *Plant Physiology* **157**: 854–865. DOI 10.1104/pp.111.179028.

Zhao Y, Ma RY, Xu DL, Bi HH, Xia ZL, Peng HR (2019). *Triticum aestivum* genome-wide identification and analysis of the AP2 transcription factor gene family in wheat (L.). *Frontiers in Plant Science* **10**: 109. DOI 10.3389/fpls.2019.01286.

TABLE S1

The 87 AP2/ERF gene-coding protein sequence information in this study

Gene name	Gene Locus	Gene name	Gene Locus	Gene name	Gene Locus
SlAP2-01	Solyc09g089910.1.1	AtAP2-01	AT2G33710.1	StAP2-01	PGSC0003DMP400044184
SlAP2-02	Solyc09g066350.1.1	AtAP2-02	AT2G28550.3	StAP2-02	PGSC0003DMP400044183
SlAP2-03	Solyc10g084340.2.1	AtAP2-03	AT2G20880.1	StAP2-03	PGSC0003DMP400007140
SlAP2-04	Solyc04g077490.3.1	AtAP2-04	AT4G39780.1	StAP2-04	PGSC0003DMP400007138
SlAP2-05	Solyc04g009450.1.1	AtAP2-05	AT4G36920.2	StAP2-05	PGSC0003DMP400024338
SlAP2-06	Solyc02g067020.1.1	AtAP2-06	AT4G37750.1	StAP2-06	PGSC0003DMP400047816
SlAP2-07	Solyc02g093150.3.1	AtAP2-07	AT1G79700.1	StAP2-07	PGSC0003DMP400039351
SlAP2-08	Solyc02g092050.3.1	AtAP2-08	AT1G78080.1	StAP2-08	PGSC0003DMP400029795
SlAP2-09	Solyc02g077360.1.1	AtAP2-09	AT1G43160.1	StAP2-09	PGSC0003DMP400047100
SlAP2-10	Solyc02g064960.3.1	AtAP2-10	AT1G75490.1	StAP2-10	PGSC0003DMP400024296
SlAP2-11	Solyc02g077370.1.1	AtAP2-11	AT1G04370.1	StAP2-11	PGSC0003DMP400041406
SlAP2-12	Solyc11g061750.2.1	AtAP2-12	AT1G51190.1	StAP2-12	PGSC0003DMP400059420
SlAP2-13	Solyc11g008560.2.1	AtAP2-13	AT1G72570.1	StAP2-13	PGSC0003DMP400048525
SlAP2-14	Solyc11g010710.2.1	AtAP2-14	AT1G22190.1	StAP2-14	PGSC0003DMP400030113
SlAP2-15	Solyc11g006050.1.1	AtAP2-15	AT1G71130.1	StAP2-15	PGSC0003DMP400063022
SlAP2-16	Solyc12g056980.1.1	AtAP2-16	AT1G16060.1	StAP2-16	PGSC0003DMP400011494
SlAP2-17	Solyc12g042210.2.1	AtAP2-17	AT1G36060.1	StAP2-17	PGSC0003DMP400023722
SlAP2-18	Solyc01g090300.2.1	AtAP2-18	AT3G54320.1	StAP2-18	PGSC0003DMP400028388
SlAP2-19	Solyc01g091760.2.1	AtAP2-19	AT3G20840.1	StAP2-19	PGSC0003DMP400044038
SlAP2-20	Solyc01g090310.2.1	AtAP2-20	AT3G23230.1	StAP2-20	PGSC0003DMP400024044
SlAP2-21	Solyc08g007830.1.1	AtAP2-21	AT5G47230.1	StAP2-21	PGSC0003DMP400028292
SlAP2-22	Solyc08g078190.1.1	AtAP2-22	AT5G17430.1	StAP2-22	PGSC0003DMP400021525
SlAP2-23	Solyc07g018290.3.1	AtAP2-23	AT5G43410.1	StAP2-23	PGSC0003DMP400004276
SlAP2-24	Solyc07g053740.1.1	AtAP2-24	AT5G67180.1	StAP2-24	PGSC0003DMP400041754
SlAP2-25	Solyc03g005500.1.1	AtAP2-25	AT5G57390.1	StAP2-25	PGSC0003DMP400021527
SlAP2-26	Solyc03g044300.3.1	AtAP2-26	AT5G50080.2	StAP2-26	PGSC0003DMP400021524
SlAP2-27	Solyc03g093610.1.1	AtAP2-27	AT5G51990.1	StAP2-27	PGSC0003DMP400008747
SlAP2-28	Solyc03g006320.1.1	AtAP2-28	AT5G10510.2	StAP2-28	PGSC0003DMP400019514
SlAP2-29	Solyc03g117720.3.1	AtAP2-29	AT5G65510.3		
		AtAP2-30	AT5G13330.1		

Supplementary file

>SlAP2-01

MELKLNMEQDPSNERRTNGGGGEIKYRGVRRRPWGKFAAEIRDSARQGARVWLGTFNTAE

EAARAYDRAAYSMRGHLAILNFPEEYNLPSSSSHFYSAGSYSSSSMASSSSSSSSRQVLE

FEYLDDKLLEELLDCDEEPNKRK*

>SlAP2-02

MENEDQRGRRENKYRGIRRRPWGKYAAEIRDPNRNGARLWLGTFETAEEAARAYDRAAFS

LRGHQAILNFPNDGHYHINNNTNVPLGVGPSYNTPSTMGNNMNMSSSSSASASASASAFSDD

RHDHGEKVEFEYLDNNLLDELLASQVHRHDNKRPKF*

>SlAP2-03

MFDLNLSFDDVSVEPISVLASEKVDEMLFNQMENSGTSNSSIVNMETSSTAADDEYFSPD

HHRDGYAFDILKSDREAGDRGGFVVKELFPLVNGEVGLDLSANYGDLNEQRIGKTPQQIL

QAKKSRRGPRSRSSQYRGVTFYRRTGRWESHIWDCGKQVYLGGFDTAHAAARAYDRAAIK

FRGLDADINFNISDYEEDLKQMKNFSKEEFVHILRRQSNGFSRGSSKYRGVTLHKCGRWE

ARMGQLLGKKYIYLGLFDNEIEAARAYDKAAIKCNGREAVTNFEPNTYEEALSSEADTGG

TDHNIDLNLGISLSSYAENQHGNTSQIGNFQCQHGSNGLPEHHGEVRTPASTTLKSKLLL

HGEHTLTQNPLHWNGSNGKGSAIEKGSEVDTSLKWMGHDQNTYDGSPTLPLFSTAASSGF

GNSANTAPSAASHQLHFGSGALPFPHSPSLTNMNLSHHYCRS*

>SlAP2-04

MKMKSMNDDSRSSNNNISNNNSAAANTNWLGFSL
TPHMKMEVTNSSDTQHQHQFAQSFYL
SSSPPPPMNVSTTSALCYENNPFHSTLSVMPLKSDGS
LCIMEALSRSHADAMVQSSSPKL
EDFLGGASMGNSQYGSHDREAMALSLDSLYYHQNE
EEIQVHSHHPYYSPMHCHGMYHQET
LLEETKPTQISNCEAQMTGNELKSWGQYAEQHAID
QHINATCSMVAAAAAAAASGGGGGTS
GCNELQSLSLSMNPGSQSSCVTPRQISPSGLECVAVE
SKKRASGKVAQKQPVHRKSIDTF
GQRTSQYRGVTRHRWTGRYEAHLWDNSCKKEGQ
TRKGRQVYLGGYDMEEKAARAYDLAAL
KYWGPSTHINFPLENYQKELDDMKNMTRQEYVAH
LRRKSSGFSRGASIYRGVTRHHQHGR
WQARIGRVAGNKDLYLGTFSTQEEAAEAYDVAAI
KFRGVNAVTNFDISRYDVEKIMASNT
LPAGELAKRTKERESIEYNNNTGGVGGQKNEECVD
NNNNGTITDWKMVLYQTSNPSLGSS
YRNPTSFSMALQDLIGIDSMTNSNNHHHATILDHE
QNKIGNPFSNASSLVTSLGSSREAS
PDKSATAASLVFAKPTKFAVPTATSVNACIPSAQLR
PIPVSMAHLPVFAALNDA*
>SlAP2-05
MESAGANAEENRSSVSAHHMYDKMPSSNSTNKSVP
AAERSDSIVLGLQNAANFSGVRKRK
RGEYVAEKWDPITKKRVCLGTFNTAEEASEAYLAK
QCEFSTARHMLDEMPSSNSTNKTAN
AAAGSTDADHGKKPLIQEISFSHKSMLLGAEWHNM
GKLRKKTFKKKQVCKGFNSNMHASN
KPINVMNLESCKQTSCSLSKSADRCNTMKRNPKAK
SGLLGIRKQKNGRYSAEIRDPIKHK
QVWLGTFNTIEEASQAYLSKKSEFDMLTQGYKEDK
PNNCDQTKPESSVGASLSSGGRNKK
MDSPKTTRIIGVRKNQWGKYTSEIINPISKKKIWLG
TFGTYEEASLAYQSKKLEFREIVK
AKKKCSKKIHLNAREKQVGKEKIAVNCEGFQPESV
AGETFHANLYMKVGVQRSKEAELQS
NMPVDFSTGEKQAGKEKLVVNCEAFQSESAAEVTF
HANQYMKFDVQRSKETKLQSNMPVD
FSTAEKQAEKEKLVVNFEGFLPESAAEETFHANQY
MKFDVQRSKESELQSKMPVDFSTWE
KQGSQEDTDSSMGKWVQLPDGREVKFSLELGVPII
DKYGSLLGDFCGFDNLWIYDESIY*
>SlAP2-06
MGRKRKAEGLMKIQRPRKKFVGVRQRPSGRWVAE
IKDTIQKIRVWLGTFDTAEEAARAYD
EAACLLRGPNTRTNFWTSSSPSSNSALPQKITKLLLSR
LREQNKSAAAAAADSSSSTTSL
AEIDHQQQKQQEKIGNRVVDFSDSLYTDYLNYPED
NVTENNVIAPITRELTSIVQNQEIN
FQPVNNYEIIEIGEAINIDVEDIESDIDFQFRFSPFDLA
EELSMDFGEETSIVSEAMRRM
NYERKFSASLYAFNGITECLKLKMKSGGVTRSDQLS
RIQNACKRNLVKERENEERNAGND
*
>SlAP2-07
MWDLNESEEGCSSPIEFEGDDEKGKRVGSVSNSSSS
AVAVDDISEEELDGERGKKKRGKI
FGFSMVGLGNGDEEQPVTRQFFPVDESEMGGVAA
ENGCPNFPRAHWVGVKFYQTETLGNT
GLARPVDMVQQQQQPIKKSRRGPRSRSSQYRGVTF
YRRTGRWESHIWDCGKQVYLGGFDT
AHAAARAYDRAAIKFRGVEADINFSLEDYESDLKQ
MTSLTKEEFVHVLRRQSTGFPRGSS
KYRGVTLHKCGRWEARMGQFLGKKYVYLGLFDTE
IEAARAYDKAAIKCNGKDAVTNFDPS
IYENELNSIESTDNVADHNLDLSLGGSSSKQGSQE
LGDNRGQNSSSIMQLDIDWQRHGLR
PEKQSALIDARRRENRYNETETLQLLSQTHLHSPA
SLKHNNNNNSQVQRFGQFMRPGDQS
HMIQMFPQQFGSSNYQIQFPSGSNGGRIGATNVRD
LSLAATSNGSSQWQSNFPPQIFVAA
AASSGFPQQIVRPQNWSSENGFHHSLMRPS*
>SlAP2-08
MKSLGNDDNSSNNNNNNWLGFSLSPHINSLEGPLT
DHHHSTQPGSSSDVQSSVPIRFSPT
HLNYPAMYYEGDNATLFSSLSAMPLKSDGSLCIMK
GFNRSQQPQGMISSTPKLEDFFGDA
AMGSHHYEGCNRGGMAHSLHNIYYNQHQDNETS
NSQDFLNHIQENTRHQQQQQNYPDFSV
FRGHEVYHSTEQGKTECSNIQVPTLTGDEMSGMNN
WVSRNYQNQTGHALEQNMIGCMADN
GAESGSVGAMTYGDFKSLSLSMSPGSQSSCVTGTQ
TQQISPTLTDCIAIETKKRGSEKAN
NQKQIVHRKSIDTFGQRTSQYRGVTRHRWTGRYEA
HLWDNSCKKEGQSRKGRQVYLGGYD
MEDKAARSYDLAALKYWGPSTHINFPLENYQQELE
EMKNMTRQEYVAHLRRKSSGFSRGA
SMYRGVTRHHQHGRWQARIGRVAGNKDLYLGTF
STQEEAAEAYDIAAIKFRGVNAVTNFD
ISRYDVERIMASNTLLAGDLARRKKETETSKEICN
QNLIANASHQGEVNIQNKEENGNAV
EWTMSLYQSSSPGMDTNSMSSTMLQGQMDESARIN
THLSNDSSLVTSLGSSREVSPAKNN
GHLMHNFAMPQSAPKLIPSPATNITSWISSAQLRP
NVPVFAAWTDA*
>SlAP2-09
MTTHHVENNNQEQDQVACEEILENVWANFIS
KNDQNSQKVTNEYCCEQYWEQLPILERLP
SLGRWISMGAETWEDILNGIIIPSHNNENSNDEST
CKDVVNVEKKEEKKKMVHYRGVRRR
PWGKYAAEIRDSSRKGARVWLGTFSTAEEAAMA
YDKAALRIRGPKAYLNFPHEMVAQAIG
ISNGPCEKEWTFSSSSQYNSRKRVSRDWNMYENLD
EINQLPMEKKIMRSMEEDLFNDLDI
LEFEDLGSDYLDSLLSSL*
>SlAP2-10
MWNLNNSPDHIKDYESEEGKGVGSISNSSSSAVEE
LYGSEEEDEFLEQNGLKGKKKKNIP
SKIFGFSMIAPPNNNNHNDDNLSSESEPPVTRQFF
PVDESEIGSGNFNDRSCRFPRAHWA
GVKFYQPEPSANSPALLGKGSELSQQVQPMKKSRR
GPRSRSSQYRGVTFYRRTGRWESHI
WDCGKQVYLGGFDTAHAAARAYDRAAIKFRGME
ADINFNLEDYEEDLKQMKNLTKEEFVH
VLRRQSTGFPRGSSKYRGVTLHKCGRWEARMGQLL
GKKYVYLGLFDTENEAARAYDKAAI
KCNGKDAVTNFDPCIYENELNSSECSNKAADHSLD
LSLGSSSSKQNSREMEDTNKNQNYP

SIQFDVDWRHQGSKPKQPSPLDMEYGRRRNVYNET
ETLQLFSQTHLHSPGSFKTNSNEMQ
RFGQYMRAGDSQMIPPQFTSSNYQVQFPSSSNRRD
TQQWQSNTVPPPHLFATAAASSGFP
QQILRHQNWPHKSGFNYSLTRPSN*
>SlAP2-11
MVPTPQSDLPLNENDSQEMVLYEVLNEANALNIPY
LPQRNQLLPRNNILRPLQCIGKKYR
GVRRRPWGKYAAEIRDSARHGARVWLGTFETAEE
AALAYDRAAFRMRGAKALLNFPSEIV
NASVSVDKLSLCSNSYTTNNNSDSSLNEVSSGTND
VFESRC*
>SlAP2-12
MNSNNWLSFPLSPTHSSLPPHLQNAQSHHFSLGLV
NETIDNPFQNQEWNLMNTQGSNEVP
KVADFLGMNKSENQSELIPYNDIQANDSDYLFQNN
HLMPSMQNALAPPPTNNYDLQENAC
NIQSLTLSMGSGKGSTSETSASPSANAATASATAE
NSNNTSIVEAAPRRTLDTFGQRTSI
YRGVTRHRWTGRYEAHLWDNSCRREGQSRKGRQ
GGYDKEEKAARAYDLAALKYWGTSTTT
NFPISNYEKELEDMKHMTRQEFVAAIRRKSSGFSR
GASMYRGVTRHHQHGRWQARIGRVA
GNKDLYLGTFTTEEEAAEAYDIAAIKFRGLNAVTN
FDMNRYDVKAILESNTLPIGGGAAK
RLKEAQALESSRKRDQEMMALNSSFQYGNSNPLQA
YPLMQQHQPSFDNSQPLLTLQNHDI
SQYNIQDSSSQFHQSYLQTQLQLQNNSHQVSLYNN
YLQNNQVFLHGLMNNNEGSSSGSYS
TGGYFGNSPGLGGMSSNSTSGNNGGGGGAHEEVA
LVKVDYDNMPSYNGWSGESSVQGSNP
GVFSMWNE*
>SlAP2-13
MASSMNNWLGFSLSPHEEVLTSQQSSHQECFDLLT
SHHHQQHHDQSTVVLPSLNVVPHDN
TPFGIFEAFNRNHHQSQDWNSHDTNYKTTSNMSM
LMGTSCNSQHLENNQEPPKLENFLGI
GEQKLQQVCNNNNNNNNNNNYYNIYCSTPNNISE
NNANNIIGLSMIKNWLRNNPNSTTTT
TTTPSSLQDDKKEGDVVGVAGGGGNGNNARSLSL
SMSMGGDGGGDNSCSSENNKQEIVSG
NNDGQISGAIVVPKKSIDTFGQRTSIYRGVTRHRW
TGRYEAHLWDNSCRREGQTRKGRQV
YLGGYDKEEKAARAYDLAALKYWGPTTTTNFPISN
YEKELEEMKHMTRQEYVASLRRKSS
GFSRGASIYRGVTRHHQHGRWQARIGRVAGNK
DLYLGTFSTQEEAAEAYDIAAIKFRGLN
AVTNFEINRYDVKSILESSTLPIGGAAKRLKDVEQ
AEIALDYQRASQENIGTTHLMNGSM
SAYGAGHCWPNITLQQAQPLSAMHYSNPYNSQQR
LWCKQEVQDHDLTSQNFQDFQLGNAH
NYLMGLDNSSSSSMEQNNVTYHQGVGYASNNNNH
HGGGGFVLPHLNGSSSYGGENDNEVK
QLQLGTHENIFGNYYHSQPATSNDSIKVSNTLYDH
QESVCNSNNWMSTAVPMALASRTST
AAICHGVAPTFTMWNDS*
>SlAP2-14
MAPDNNNWLSFSLSSMEMLNSTSSHHHQNNSNNN
NSSLLHHHHQSNMKFDLASSSADSHQ
YYFADNFFPHGWSNPKPQVMYSEAEKEVIGLNINT
DSSIFQNFVENAPPKLEDFFGGDSS
SLTQDSSSLTHIYENPNGSAVYDNNNYNEHQDFKN
ITGFQAFSTNSGSEVDNSVGTEFGT
NESCNELSYSQCVAAAVAAAAPPLLPTGGGGGALS
LAVNAAAATTTNQCSENEKAMVAVA
DSQSCKKISDTFGQRTSIYRGVTRHRWTGRYEAHL
WDNSCRREGQARKGRQVYLGGYDKE
DKAARAYDLAALKYWGPTATTNFPVTNYTKELEE
MKHMTKQEFIASLRRKSSGFSRGASI
YRGVTRHHQQGRWQARIGRVAGNKDLYLGTFATE
EEAAEAYDIAAIKFRGVNAVTNFEMN
RYDVEAIMKSSLPVGGTAKRLKLSLESEQKSSSSN
NNNQQQQMTQCNSSNNINFGAMPPV
SAIPCGVPFDNTAPAFYHHNFFQYLHPGNASVPDA
SNATTTMAAAMPLLPSAADLFIWPH
QSY*
>SlAP2-15
MDSSSSSSSYSYSQYHDHDLGFNENDTEDMLLLNVL
SEAKEDSSSNNIHEETNEGSSKEV
ATSYRGVRRRPWGKYAAEIRDSTRNSVRVWLGTFD
TAEEAALAYDQAAYALHGRVAVLNY
PVEMVYRSLMEIDCRFEDGCSPVLALKKWHSMMK
KGNNKSSSSNKKKEEKEMSYENVLVL
EDLGADYLEEILMLSESNS*
>SlAP2-16
MATSTMDFWTTTLLDLNSSNSGGELMEALAPFIKS
ASSPSPSPSVSPSFDLQSSSLSTSF
LYESFSSTSQPNMSSIGLNQAQIYLSQQVMPAVTFQ
NNNQYASYLGPKPVSMKQTGSPPK
PPKLYRGVRQRHWGKWVAEIRLPKNRTRLWLGTF
DTAEEAALAYDKAAYKLRGEFARLNF
PHLRHNGSLIGSEFGEYKPLHSSVNAKLQAICQDLA
QGKSIDTKKKRKVSSKAMMVEVEE
KEYKKSKTTAEAGSESDGSGSGSGSGSGSGSSPISE
YTFDSIWDMCSENYVLHKDPSQEI
FNWASLL*
>SlAP2-17
MVDRRYGKRPFSSNEWEEKEDINF1PIYSARSQHDM
SAMVSVLSQVISNTNNTTNISSSSS
MHEIDPLTLPQPTTNQIHEQGNQQRKRHYRGVRQ
RPWGKWAAEIRDPKKAARVWLGTFDT
AEAAALAYDEAALRFKGNKAKLNFPERVQSGNTQY
LTTTHQQQQQHYAFNNNNNNIPQMV
TQPNLYQQHFPNVHHYAQLLRDGSNNNIDNMMN
FGVSDQSSFYHHHQQGNFISPNTSLEL
QQQQQQASYYHNQQEDFLRYTMDFGNSSNYSTGP
PSESNWMDFEPKK*
>SlAP2-18
MYQQKSTTISDSDLSVLENIKFHLVNDSDFSQILSMF
DPINISHADIINSPNSSYGSSTS
AAEISWGDMITNIDSPWQCIDKLEHEEAPKEEPLV
ARGVHAPGDWNRYRGVRRRPWGKFA
AEIRDPDRKGARLWLGTYGTPEDAALAYDQAAYKI
RGSKARLNFPHLIGSDISEPVRVAP
RKRCHSSQSSLVEYTSSKKRKL*
>SlAP2-19
MATDLENTKTTITTNTNTTNSSSTEAKKSSITRKF
VGVRQRPSGRWVAEIKDSSQRVRLW

LGTYDTPEEAAHAYDEAARALRGENARTNFASTTP
NSDMDQSNILHSKNGLSFSSIKAKL
SKNLQSIMARNSENKSSSSKSITRVSDHFTFARIF
HFKNNYDQPYQNHRHVDMNKVVQPS
IRVLPHDVTDNNNNNDSSWENSSSVSDCSSEWAAF
RQLGLDCDNYGSDGSEYFVGSDPLM
AGWMSSPDIMSSTSTNEGSSRSKRFKVSSSVVVPP
TFTESPLHDAQNYVPF*
>SlAP2-20
MNPSDFSLLQSIQHHLLNDSDFPNIFSAIDSNNTP
TDFTQNNFDYGELTPLINSSTTLQA
NEKSEVEESETVVASVKHAPKDWKRYRGVRRRPW
GKFAAEIRDPDKKNARLWLGTYETPE
DAALAYDQAAFKIRGSKARLNFPHLIGSGVPEPAR
VNPRRRSHSPESSCENGTPRKLYFI
*
>SlAP2-21
MNSDHFRVQDEGVFLLASSQPKKRAGRKKFKETRH
PVYRGVRRRDNNKWVCEVREPSEQK
RIWLGTYPTPEMAARAHDVAALALRGNLATLNFA
DSSWRLTVPISKDPEELRKAAIQAAE
ELVGVGYMNEEINYIEENIAIGGANGSDNSNLVGN
EVNNNNNIQNMEMGNISCYNNWGDN
NEMLEIEGSTWQEKMSEGHLFSPTPHYDSCFSWDD
VESDGEVSLWSYNM*
>SlAP2-22
MEYVDEFFALEKIRTHLLGEFSPRTLKFGAELTSSKC
SDTVTESCDSESASSGSLPFDFF
FDIEADLFQFGAGVSNSSSESANSVYNRTDFVKIESE
PSISSSDYEENQSKRFQFKSEPQ
VFIDLTSPKSRKLCSASDRKPSLKIDLPPVKKYEWID
FGNSAQSNPIVSVPVKQIREAEE
KRHYRGVRQRPWGKYAAEIRDPRRRGSRVWLGTF
DTAIEAAKAYDKAAFTMRGSKAILNF
PLEVGKTLSYSSSAVEGGVKRRREAVETEEEKVVK
KCKEEESCPLTPSSWNFVFEQNCNG
MFNLPPLSPLSPHLSCSQLIQI*
>SlAP2-23
MDMNSSSCEKWLGFSLSNINNLSPSTTNSSQLSLFQA
FNSNPNTNCEVVRQNENGEPKVE
NMLGSCCDSNTPTSCEYGSELKRIAATFLPTTFSDD
HKEKLPLIVTQPKKPHESFGQRTS
IYRGVTRHRWTGRYEAHLWDNSCRRQGQSRKGRQ
VYLGGYDKEEKAARAYDLAALKYWGP
TTTTNFPVSNYERELEEMKNMTRQEFVASLRRKSS
GFSRGASIYRGVTRHHQHGRWQARI
GRVAGNKDLYLGTFSTQEEAAEAYDIAAIKFRGLN
AVTNFDISRYDVKSIASSNLPVGGM
SNKSKSSSDETTKHVENRPNLRDQDQDRDISSAPN
YTSKNLAMVNLGFGLPIKQDAFDFW
SSVGYNNNRSKNGLLFQGTTMNAQGTTFFSNEVN
NSSVMLNEQGYDNHDQQQQSGGSSCE
INMALNSNIATTSTTIDSTNFGNWMTPSLHSFQS
STKNDLGPYHTPIFGME*
>SlAP2-24
MAVKDKAVKGGNVKVNHGVKEVHYRGVRKRPW
GRYAAEIRDPGKKSRVWLGTFDTAEEAA
KAYDAAAREFRGPKAKTNFPFPAEMNNVGNNNSQ
SPCGSSTVESSSGETVVHAPNTRHAP
LELDLTRRLGAAAEGGRGGVGYPILHQQPTVAVL
PNGQPVLLFDSMWRPGVVSRPYQVVP
ATMEFAGVGAGVVTSVSDSSSVVEEKHYGKKGLD
LDLNLAPPMEV*
>SlAP2-25
MDEVKYRGVRKRPWGKYAAEIRDTNRQGGVRVW
LGTFSSAEEAARAYDKAAYNMRGHLAI
LNFPEEYNLPRSSSHFYNNSSSMPSSSSSSNVHARND
QQGNRQVLELEYLDDHVLEELLD
CDQHANNRK*
>SlAP2-26
MWNLNDSPDQTMEYESDEGITVRSESNSISSALLVV
EDGNSSEEDGEKGKKKKSNNTPGK
IFGFSIKDHNLESPVVTRQFFPVDNESTNFPRAQWA
GIKFCESEPPLVNGLVGNKIDVLQ
QQPIKKSRRGPRSRSSQYRGVTFYRRTGRWESHIWD
CGKQVYLGGFDTAHAAARAYDRAA
IKFRGAEADINFTSKDYEDDLKQMSNLTKEEFVHVL
RRQSTGFPRGSSKYRGVTLHKCGR
WEARMGQFLGKKYVYLGLFDTEVEAARAYDKAAI
KCNGKDAVTNFDRSIYENELNSTECT
DNATDHNLDLSLGGSSQEMGDNRGQNSSSNLQLD
GHWGHQGSSRHNNKVQTPSSNNVGQI
GGTSNFQNNYEIIATAAASSGFPQQIIRTQNNGFHH
YFMRP*
>SlAP2-27
MYQLPTSTELTFFPAEFPVYCRSSSFSSLMPCLTESW
GDLPLKVNDSEDMVIYGFLQDAF
SIGWTPSNLTSEEVKLEPREEIEPAMSTSVSPPTVAP
AALQPKGRHYRGVRQRPWGKFAA
EIRDPAKNGARVWLGTYESAEEAALAYDKAAFRM
RGTKALLNFPHRIGLNEPEPVRVTVK
RRLSESASSSVSSASESGSPKRRRKGVAAKQAELEVE
SRGPNVMKVGCQMEQFPVGEQLL
VS*
>SlAP2-28
MAPKPKCTTAATAATAAVVKEVHYRGVRKRPWG
RYAAEIRDPGKKCRVWLGTFDTAEEAA
RAYDKAAIEFRGAKAKTNFQMQQQPDDVIRSPSD
TSTVESSSAAVAVAKAPAMVESIPLD
LSLGSSSSAVGIFSSGVGKFLFQNSPPMYYFQGAGVI
RNGGGGGGGSGDGGGAGASGAGG
MMKSESDSSTVIDFMGNNFKPKAKFDLNLLPTPE
DM*
>SlAP2-29
MAKTSKSNTTSTSSSSSSNKCDSKAKRSNKIDGNAI
GKVKRTRKSVPRDSPPQRSSIYRG
VTRHRWTGRYEAHLWDKNCWNETQNKKGRQVY
LGAYDDEEAAAHAYDLAALKYWGQDTML
NFPIMTYENELKEMEGQSKEEYIGSLRRKSSGFSRGV
SKYRGVARHHHNGRWEARIGRVF
GNKYLYLGTYATQEEAATAYDMAAIEYRGLNAVT
NFDLSRYIKWLRPSDQTNNDNTIINP
EPNPNPNPNDIHLMPNTKDDTNFTQQQQQISGCD
VTVAALPHPGGGAATSSAALELLLQS
TKLKEMLERRSEVIECPETPPEPDRPRRSFPDDIQTY
FDCQEPSSFIEEHDIIFGDLDSL
TLPMFQCELIN*
>AtAP2-01

MHSGKRPLSPESMAGNREEKKELCCCSTLSESDVSD
FVSELTGQPIPSSIDDQSSSLTLQ
EKSNSRQRNYRGVRQRPWGKWAAEIRDPNKAARV
WLGTFDTAEEAALAYDKAAFEFRGHK
AKLNFPEHIRVNPTQLYPSPATSHDRIIVTPPSPPPPI
APDILLDQYGHFQSRSSDSSAN
LSMNMLSSSSSSLNHQGLRPNLEDGENVKNISIHKR
RK*
>AtAP2-02
MLDLNLNADSPESTQYGGDSYLDRQTSDNSAGNRV
EESGTSTSSVINADGDEDSCSTRAF
TLSFDILKVGSSSGGDESPAASASVTKEFFPVSGDCG
HLRDVEGSSSSRNWIDLSFDRIG
DGETKLVTPVPTPAPVPAQVKKSRRGPRSRSSQYRG
VTFYRRTGRWESHIWDCGKQVYLG
GFDTAHAAARAYDRAAIKFRGVDADINFTLGDYEE
DMKQVQNLSKEEFVHILRRQSTGFS
RGSSKYRGVTLHKCGRWEARMGQFLGKKYIYLGLF
DSEVEAARAYDKAAINTNGREAVTN
FEMSSYQNEINSESNNSEIDLNLGISLSTGNAPKQNG
RLFHFPSNTYETQRGVSLRIDNE
YMGKPVNTPLPYGSSDHRLYWNGACPSYNNPAEG
RATEKRSEAEGMMSNWGWQRPGQTSA
VRPQPPGPQPPPLFSVAAASSGFSHFRPQPPNDNA
TRGYFYPHP*
>AtAP2-03
MATAKNKGKSIRVLGTSEAEKKDEMELEEEFQFSSG
KYKDSGPGSDMWLGDASSTSPRSL
RKTRTFDRHNPYLVSSYATPQPPTTTTCSVSFPFYL
PPAIQNQQRFLHPNDPSGQRQQQM
ISFDPQQQVQPYVAQQQQQQQHLLQYWRDILKLS
PSGRMMMMNMLRQESDLPLTRPPVQP
FSATKLYRGVRQRHWGKWVAEIRKPRNRTRLWLG
TFDTAEEAAMAYDREAFKLRGETARL
NFPELFLNKQEPTPVHQKQCETGTTSEDSSRRGEDD
SSTALAVGGVSEETGWAEAWFNAI
PEEWGPGSPLWDDYHFPISNHKDDLDATQNSSSDTI*
>AtAP2-04
MAAIDMFNSNTDPFQEELMKALQPYTTNTDSSSPT
YSNTVFGFNQTTSLGLNQLTPYQIH
QIQNQLNQRRNIISPNLAPKPVPMKNMTAQKLYR
GVRQRHWGKWVAEIRLPKNRTRLWLG
TFDTAEEAAMAYDLAAYKLRGEFARLNFPQFRHE
DGYYGGGSCFNPLHSSVDAKLQEICQ
SLRKTEDIDLPCSETELFPPKTEYQESEYGFLRSDENS
FSDESHVESSSPESGITTFLDF
SDSGFDEIGSFGLEKFPSVEIDWDAISKLSES*
>AtAP2-05
MWDLNDAPHQTQREEESEEFCYSSPSKRVGSFSNSS
SSAVVIEDGSDDDELNRVRPNNPL
VTHQFFPEMDSNGGGVASGFPRAHWFGVKFCQSD
LATGSSAGKATNVAAAVVEPAQPLKK
SRRGPRSRSSQYRGVTFYRRTGRWESHIWDCGKQV
YLGGFDTAHAAARAYDRAAIKFRGV
EADINFNIDDYDDDLKQMTNLTKEEFVHVLRRQST
GFPRGSSKYRGVTLHKCGRWEARMG
QFLGKKYVYLGLFDTEVEAARAYDKAAIKCNGKD
AVTNFDPSIYDEELNAESSGNPTTPQ
DHNLDLSLGNSANSKHKSQDMRLRMNQQQQDS
LHSNEVLGLGQTGMLNHTPNSNHQFPGS
SNIGSGGGFSLFPAAENHRFDGRASTNQVLTNAA
ASSGFSPHHHNQIFNSTSTPHQNWLQ
TNGFQPPLMRPS*
>AtAP2-06
MKSFCDNDDNNHSNTTNLLGFSLSSNMMKMGGRG
GREAIYSSSTSSAATSSSSVPPQLVV
GDNTSNFGVCYGSNPNGGIYSHMSVMPLRSDGSL
CLMEALNRSSHSNHHQDSSPKVEDFF
GTHHNNTSHKEAMDLSLDSLFYNTTHEPNTTTNF
QEFFSFPQTRNHEEETRNYGNDPSLT
HGGSFNVGVYGEFQQSLSLSMSPGSQSSCITGSH
HHQQNQNQNHQSQNHQQISEALVETS
VGFETTTMAAAKKKRGQEDVVVVGQKQIVHRKSI
DTFGQRTSQYRGVTRHRWTGRYEAHL
WDNSFKKEGHSRKGRQVYLGGYDMEEKAARAYDL
AALKYWGPSTHTNFSAENYQKEIEDM
KNMTRQEYVAHLRRKSSGFSRGASIYRGVTRHHQ
HGRWQARIGRVAGNKDLYLGTFGTQE
EAAEAYDVAAIKFRGTNAVTNFDITRYDVDRIMS
SNTLLSGELARRNNNSIVVRNTEDQT
ALNAVVEGGSNKEVSTPERLLSFPAIFALPQVNQ
KMFGSNMGGNMSPWTSNPNAELKTVA
LTLPQMPVFAAWADS*
>AtAP2-07
MAKVSGRSKKTIVDDEISDKTASASESASIALTS
KRKRKSPPRNAPLQRSSPYRGVTRHR
WTGRYEAHLWDKNSWNDTQTKKGRQVYLGAYD
EEEAAARAYDLAALKYWGRDTLLNFPLP
SYDEDVKEMEGQSKEEYIGSLRRKSSGFSRGVSK
YRGVARHHHNGRWEARIGRVFATQEE
AAIAYDIAAIEYRGLNAVTNFDVSRYLNPNAAAD
KADSDSKPIRSPSREPESSDDNKSPK
SEEVIEPSTSPEVIPTRRSFPDDIQTYFGCQDSG
KLATEEDVIFDCFNSYINPGFYNEFD
YGP*
>AtAP2-08
MAAAMNLYTCSRSFQDSGGELMDALVPFIKSVSD
SPSSSSAASASASAFLHPSAFSLPPLPG
YYPDSTFLTQPFSYGSDLQQTGSLIGLNNLSSSQ
IHQIQSQIHHPLPPTHHNNNNSFSNL
LSPKPLLMKQSGVAGSCFAYGSGVPSKPTKLYRG
VRQRHWGKWVAEIRLPRNRTRLWLGT
FDTAEEAALAYDKAAYKLRGDFARLNFPNLRHNG
SHIGGDFGEYKPLHSSVDAKLEAICK
SMAETQKQDKSTKSSKKREKKVSSPDLSEKVKAE
ENSVSIGGSPPVTEFEESTAGSSPLS
DLTFADPEEPPQWNETFSLEKYPSYEIDWDSILA*
>AtAP2-09
MVSMLTNVVSGETEPSASATWTMGHKREREEFSL
PPQPLITGSAVTKECESSMSLERPKK
YRGVRQRPWGKWAAEIRDPHKATRVWLGTFETAE
AAARAYDAAALRFRGSKAKLNFPENV
GTQTIQRNSHFLQNSMQPSLTYIDQCPTLLSYSR
CMEQQQPLVGMLQPTEEENHFFEKPW
TEYDQYNYSSFG*
>AtAP2-10
MSSIEPKVMMVGANKKQRTVQASSRKGCMRGKGG
PDNASCTYKGVRQRTWGKWVAEIREP
NRGARLWLGTFDTSREAALAYDSAARKLYGPEAH
LNLPESLRSYPKTASSPASQTTPSSN

TGGKSSSDSESPCSSNEMSSCGRVTEEISWEHIN
VDLPVMDDSSIWEEATMSLGFPWVHE
GDNDISRFDTCISGGYSNWDSFHSPL*
>AtAP2-11
MDQGGRSSGSGGGGAEQGKYRGVRRRPWGKYAA
EIRDSRKHGERVWLGTFDTAEDAARAY
DRAAYSMRGKAAILNFPHEYNMGTGSSSTAANSS
SSSQQVFEFEYLDDSVLDELLEYGEN
YNKTHNINMGKRQ*
>AtAP2-12
MNSNNWLAFPLSPTHSSLPPHIHSSQNSHFNLGL
VNDNIDNPFQNQGWNMINPHGGGGEG
GEVPKVADFLGVSKSGDHHTDHNLVPYNDIHQTN
ASDYYFQTNSLLPTVVTCASNAPNNY
ELQESAHNLQSLTLSMGSTGAAAAEVATVKASPA
ETSADNSSSTTNTSGGAIVEATPRRT
LETFGQRTSIYRGVTRHRWTGRYEAHLWDNSCRR
EGQSRKGRQVYLGGYDKEEKAARAYD
LAALKYWGPSTTTNFPITNYEKEVEEMKNMTRQE
FVASIRRKSSGFSRGASMYRGVTRHH
QHGRWQARIGRVAGNKDLYLGTFSTEEEAAEAYD
IAAIKFRGLNAVTNFEINRYDVKAIL
ESNTLPIGGGAAKRLKEAQALESSRKREEMIALG
SNFHQYGAASGSSSVASSSRLQLQPY
PLSIQQPFEHLHHHQPLLTLQNNNDISQYHDSFS
YIQTQLHLHQQQTNNYLQSSSHTSQL
YNAYLQSNPGLLHGFVSDNNNTSGFLGNNGIGIG
SSSTVGSSAEEEFPAVKVDYDMPPSG
GATGYGGWNSGESAQGSNPGGVFTMWNE*
>AtAP2-13
MKKWLGFSLTPPLRICNSEEEELRHDGSDVWRYD
INFDHHHHDEDVPKVEDLLSNSHQTE
YPINHNQTNVNCTTVVNRLNPPGYLLHDQTVVTP
HYPNLDPNLSNDYGGFERVGSVSVFK
SWLEQGTPAFPLSSHYVTEEAGTSNNISHFSNEE
TGYNTNGSMLSLALSHGACSDLINES
NVSARVEEPVKVDEKRKRLVVKPQVKESVPRKSV
DSYGQRTSQYRGVTRHRWTGRYEAHL
WDNSCKKEGQTRRGRQVYLGGYDEEEKAARAYDL
AALKYWGPTTHLNFPLSNYEKEIEEL
NNMNRQEFVAMLRRNSSGFSRGASVYRGVTRHHQ
HGRWQARIGRVAGNKDLYLGTFSTQE
EAAEAYDIAAIKFRGLNAVTNFDINRYDVKRICS
SSTIVDSDQAKHSPTSSGAGH*
>AtAP2-14
MTTSMDFYSNKTFQQSDPFGGELMEALLPF
IKSPSNDSSAFAFSLPAPISYGSDLHSFSH
HLSPKPVSMKQTGTSAAKPTKLYRGVRQRHWGKW
VAEIRLPRNRTRLWLGTFDTAEEAAL
AYDKAAYKLRGDFARLNFPDLRHNDEYQPLQSSV
DAKLEAICQNLAETTQKQVRSTKKSS
SRKRSSTVAVKLPEEDYSSAGSSPLLTESYGSGG
SSSPLSELTFGDTEEEIQPPWNENAL
EKYPSYEIDWDSILQCSSLVN*
>AtAP2-15
MKRIIRISFTDAEATDSSSDEDTEERGGASQTRR
RGKRLVKEIVIDPSDSADKLDVCKTR
FKIRIPAEFLKTAKTEKKYRGVRQRPWGKWVAEI
RCGRGACKGRRDRLWLGTFNTAEEAA
LAYDNASIKLIGPHAPTNFGLPAENQEDKTVIGA
SEVARGA*
>AtAP2-16
MFIAVEVSPVMEDITRQSKKTSVENETGDDQSAT
SVVLKAKRKRRSQPRDAPPQRSSVHR
GVTRHRWTGRYEAHLWDKNSWNETQTKKGRQV
YLGAYDEEDAAARAYDLAALKYWGRDTI
LNFPLCNYEEDIKEMESQSKEEYIGSLRRKSSGF
SRGVSKYRGVAKHHHNGRWEARIGRV
FGNKYLYLGTYATQEEAAIAYDIAAIEYRGLNAV
TNFDISRYLKLPVPENPIDTANNLLE
SPHSDLSPFIKPNHESDLSQSQSSSEDNDDRKTK
LLKSSPLVAEEVIGPSTPPEIAPPRR
SFPEDIQTYFGCQNSGKLTAEEDDVIFGDLDSFL
TPDFYSELNDC*
>AtAP2-17
MADLFGGGHGGELMEALQPFYKSASTSASN
PAFASSNDAFASAPNDLFSSSSYYNPHASL
FPSHSTTSYPDIYSGSMTYPSSFGSDLQQPENYQ
SQFHYQNTITYTHQDNNTCMLNFIEP
SQPGFMTQPGPSSGSVSKPAKLYRGVRQRHWGKW
VAEIRLPRNRTRLWLGTFDTAEEAAL
AYDRAAFKLRGDSARLNFPALRYQTGSSPSDTGE
YGPIQAAVDAKLEAILAEPKNQPGKT
ERTSRKRAKAAASSAEQPSAPQQHSGSGESDGSG
SPTSDVMVQEMCQEPEMPWNENFMLG
KCPSYEIDWASILS*
>AtAP2-18
MKKRLTTSTCSSSPSSSVSSSTTTSSPIQSEAPR
PKRAKRAKKSSPSGDKSHNPTSPAST
RRSSIYRGVTRHRWTGRFEAHLWDKSSWNSIQNK
KGKQVYLGAYDSEEAAAHTYDLAALK
YWGPDTILNFPAETYTKELEEMQRVTKEEYLASL
RRQSSGFSRGVSKYRGVARHHHNGRW
EARIGRVFGNKYLYLGTYNTQEEAAAAYDMAAIE
YRGANAVTNFDISNYIDRLKKKGVFP
FPVNQANHQEGILVEAKQEVETREAKEEPREEVK
QQYVEEPPQEEEEKEEEKAEQQEAEI
VGYSEEAAVVNCCIDSSTIMEMDRCGDNNELAWN
FCMMDTGFSPFLTDQNLANENPIEYP
ELFNELAFEDNIDFMFDDGKHECLNLENLDCCVV
GRESPPSSSSPLSCLSTDSASSTTTT
TTSVSCNYLFQGLFVGSE*
>AtAP2-19
MNSNNWLGFPLSPNNSSLPPHEYNLGLVSDHMDN
PFQTQEWNMINPHGGGGDEGGEVPKV
ADFLGVSKPDENQSNHLVAYNDSDYYFHTNSLMP
SVQSNDVVVAACDSNTPNNSSYHELQ
ESAHNLQSLTLSMGTTAGNNVVDKASPSETTGDN
ASGGALAVVETATPRRALDTFGQRTS
IYRGVTRHRWTGRYEAHLWDNSCRREGQSRKGRQ
VYLGGYDKEDKAARSYDLAALKYWGP
STTTNFPITNYEKEVEEMKHMTRQEFVAAIRRKS
SGFSRGASMYRGVTRHHQHGRWQARI
GRVAGNKDLYLGTFSTEEEAAEAYDIAAIKFRGL
NAVTNFEINRYDVKAILESSTLPIGG
GAAKRLKEAQALESSRKREAEMIALGSSFQYGGG
SSTGSGSTSSRLQLQPYPLSIQQPLE
PFLSLQNNDISHYNNNNAHDSSSFNHHSYIQTQL
HLHQQTNNYLQQQSSQNSQQLYNAYL

HSNPALLHGLVSTSIVDNNNNNGGSSGSYNTAAF
LGNHGIGIGSSSTVGSTEEFPTVKTD
YDMPSSDGTGGYSGWTSESVQGSNPGGVFTMWNE*
>AtAP2-20
MESSNRSSNNQSQDDKQARFRGVRRRPWGKFAAE
IRDPSRNGARLWLGTFETAEEAARAY
DRAAFNLRGHLAILNFPNEYYPRMDDYSLRPPYA
SSSSSSSSGSTSTNVSRQNQREVFEF
EYLDDKVLEELLDSEERKR*
>AtAP2-21
MATPNEVSALWFIEKHLLDEASPVATDPWMKHES
SSATESSSDSSSIIFGSSSSSFAPID
FSESVCKPEIIDLDTPRSMEFLSIPFEFDSEVSV
SDFDFKPSNQNQNQFEPELKSQIRKP
PLKISLPAKTEWIQFAAENTKPEVTKPVSEEEKK
HYRGVRQRPWGKFAAEIRDPNKRGSR
VWLGTFDTAIEAARAYDEAAFRLRGSKAILNFPL
EVGKWKPRADEGEKKRKRDDDEKVTV
VEKVLKTEQSVDVNGGETFPFVTSNLTELCDWDL
TGFLNFPLLSPLSPHPPFGYSQLTVV
*
>AtAP2-22
MNSMNNWLGFSLSPHDQNHHRTDVDSSTTRTAV
DVAGGYCFDLAAPSDESSAVQTSFLSP
FGVTLEAFTRDNNSHSRDWDINGGACNNINNNEQ
NGPKLENFLGRTTTIYNTNETVVDGN
GDCGGGDGGGGGSLGLSMIKTWLSNHSVANAN
HQDNGNGARGLSLSMNSSTSDSNNYNNN
DDVVQEKTIVDVVETTPKKTIESFGQRTSIYRGV
TRHRWTGRYEAHLWDNSCKREGQTRK
GRQVYLGGYDKEEKAARAYDLAALKYWGTTTTTN
FPLSEYEKEVEEMKHMTRQEYVASLR
RKSSGFSRGASIYRGVTRHHQHGRWQARIGRVAG
NKDLYLGTFGTQEEAAEAYDIAAIKF
RGLSAVTNFDMNRYNVKAILESPSLPIGSSAKRL
KDVNNPVPAMMISNNVSESANNVSGW
QNTAFQHHQGMDLSLLQQQQERYVGYYNGGNL
STESTRVCFKQEEEQQHFLRNSPSHMTN
VDHHSSTSDDSVTVCGNVVSYGGYQGFAIPVGTS
VNYDPFTAAEIAYNARNHYYYAQHQQ
QQQIQQSPGGDFPVAISNNHSSNMYFHGEGGGEG
APTFSVWNDT*
>AtAP2-23
MDQGGRGVGAEHGKYRGVRRRPWGKYAAEI
RDSRKHGERVWLGTFDTAEEAARAYDQAAY
SMRGQAAILNFPHEYNMGSGVSSSTAMAGSSSAS
ASASSSSRQVFEFEYLDDSVLEELLE
EGEKPNKGKKK*
>AtAP2-24
MWNLNDSPDHHEESDSRGNPVGHVSNGMSQSAT
WLPFVLPVTRNFFPAQSMEPGVRWSGF
NSVGKSDPSGSGRPEEPEISPPIKKSRRGPRSRS
SQYRGVTFYRRTGRWESHIWDCGKQV
YLGGFDTAHAAARAYDRAAIKFRGVDADINFDIE
DYLDDLKQMGNLTKEEFMHVLRRQST
GFPRGSSKYRGVTLHKCGRWESRLGQFLNKKYVY
LGLFDTEIEAARAYDKAAIKCNGKDA
VTNFDPKVYEEEEDLSSETTRNGHNLGLSLGESS
SEEFRLKSDIASIRSRIRDEERLLGS
DLSLAMMTTTVRSEKQQSDGGGNRVVGMAASSGF
SPQPSPYRIPRTFHFSRP*
>AtAP2-25
MKNNNNKSSSSSSYDSSLSPSSSSSSHQNWLSFS
LSNNNNNFNSSSNPNLTSSTSDHHHP
HPSHLSLFQAFSTSPVERQDGSPGVSPSDATAVL
SVYPGGPKLENFLGGGASTTTTRPMQ
QVQSLGGVVFSSDLQPPLHPPSAAEIYDSELKSI
AASFLGNYSGGHSSEVSSVHKQQPNP
LAVSEASPTPKKNVESFGQRTSIYRGVTRHRWTG
RYEAHLWDNSCRREGQSRKGRQVYLG
GYDKEDKAARAYDLAALKYWGPTTTTNFPISNYE
SELEEMKHMTRQEFVASLRRKSSGFS
RGASMYRGVTRHHQHGRWQARIGRVAGNKDLYL
GTFSTQEEAAEAYDIAAIKFRGLNAVT
NFDISRYDVKSIASCNLPVGGLMPKPSPATAAAD
KTVDLSPSDSPSLTTPSLTFNVATPV
NDHGGTFYHTGIPIKPDPADHYWSNIFGFQANPK
AEMRPLANFGSDLHNPSPGYAIMPVM
QEGENNFGGSFVGSDGYNNHSAASNPVSAIPLSS
TTTMSNGNEGYGGNINWINNNISSSY
QTAKSNLSVLHTPVFGLE*
>AtAP2-26
MSAMVSALTQVVSARSQTEAEGAHSQTEAEGAHS
SSSSAGHKRGWLGIDSAPIPSSFARV
DSSHNPIEESMSKAFPEEAREKKRRYRGVRQRPW
GKWAAEIRDPHRAARVWLGTFDTAEA
AARAYDEAALRFRGNKAKLNFPEDVRILPPPPPL
LRSPADTVANKAEEDLINYWSYTKLL
QSSGQRSFLERGQEESSNIFEHSPMEQPLPPSSS
GPSSSNFPAPSLPNT*
>AtAP2-27
MNPFYSTFPDSFLSISDHRSPVSDSSECSPKLAS
SCPKKRAGRKKFRETRHPIYRGVRQR
NSGKWVCEVREPNKKSRIWLGTFPTVEMAARAHD
VAALALRGRSACLNFADSAWRLRIPE
TTCPKEIQKAASEAAMAFQNETTTEGSKTAAEAE
EAAGEGVREGERRAEEQNGGVFYMDD
EALLGMPNFFENMAEGMLLPPPEVGWNHNDFD
GVGDVSLWSFDE*
>AtAP2-28
MYFLRHETQSWSNQKPQEFFKEEAQLAAAASMAD
STILTTFVDPQSHHSQNHIPKLEDFL
GDSSSIVRYSDNSQTDTQDSSLTQIYDPRHHHNQ
TGFYSDHHDFKTMAGFQSAFSTNSGS
EVDDSASIGRTHLAGDYLGHVVESSGPELGFHGG
STGALSLGVNVNNNTNHRNDNDNHYR
GNNNGERINNNNNNDNEKTDSEKEKAVVAVETSD
CSNKKIADTFGQRTSIYRGVTRHRWT
GRYEAHLWDNSCRREGQARKGRQGGYDKEDKAA
RAYDLAALKYWNATATTNFPITNYSKE
VEEMKHMTKQEFIASLRRKSSGFSRGASIYRGVT
RHHQQGRWQARIGRVAGNKDLYLGTF
ATEEEAAEAYDIAAIKFRGINAVTNFEMNRYDVE
AIMKSALPIGGAAKRLKLSLEAAASS
EQKPILGHHQLHHFQQQQQQQQLQLQSSPNHSSI
NFALCPNSAVQSQQIIPCGIPFEAAA
LYHHHQQQQQHQQQQQQQNFFQHFPANAASDST
GSNNNSNVQGTMGLMAPNPAEFFLWPN
QSY*

>AtAP2-29
MADSTTLSTFFDHSQTQIPKLEDFLGDSFVRYSD
NQTETQDSSSLTPFYDPRHRTVAEGV
TGFFSDHHQPDFKTINSGPEIFDDSTTSNIGGTH
LSSHVVESSTTAKLGFNGDCTTTGGV
LSLGVNNTSDQPLSCNNGERGGNSNKKKTVSKKE
TSDDSKKKIVETLGQRTSIYRGVTRH
RWTGRYEAHLWDNSCRREGQARKGRQVYLGGYD
KEDRAARAYDLAALKYWGSTATTNFPV
SSYSKELEEMNHMTKQEFIASLRRKSSGFSRGAS
IYRGVTRHHQQGRWQARIGRVAGNKD
LYLGTFGKTLS*
>AtAP2-30
MVSALSRVIENPTDPPVKQELDKSDQHQPDQDQP
RRRHYRGVRQRPWGKWAAEIRDPKKA
ARVWLGTFETAEEAALAYDRAALKFKGTKAKLNF
PERVQGPTTTTTISHAPRGVSESMNS
PPPRPGPPSTTTTSWPMTYNQDILQYAQLLTSNN
EVDLSYYTSTLFSQPFSTPSSSSSSS
QQTQQQQLQQQQQQREEEEKNYGYNYYNYPRE*
>StAP2-01
MWNLNDSPDHRRDYEGKGVERSISNSSSSAVEEL
YGSEEEDEFLEQNGLKGKKKKSIPSK
IFGFSMIAPPNNNNDDNLSSESEPPVTRQFFPVD
ESEIGSGNFNDRSCRFPRAHWAGVKF
YQPEPSTAVLGKGSELSQVQVQVQVQVQPMKKSR
RGPRSRSSQYRGVTFYRRTGRWESHI
WDCGKQVYLGGFDTAHAAARAYDRAAIKFRGVEA
DINFNLEDYEEDLKQMKNLTKEEFVH
VLRRQSTGFPRGSSKYRGVTLHKCGRWEARMGQL
LGKKYVYLGLFDTENEAARAYDKAAI
NCNGKDAVTNFDPCIYENELNSSESSNKAADHSL
DLSLGSSTSKQKSREMEDTNKDQNYS
SMQFDVDWRHQGSRLKQQHSPIDMEYGRRRNV
YNETDTLQLLSQTHLHSPGSFKTNSNEM
QRFGQYMRAGDSQMIPPQFTSSNYQVQFPSSNNR
RDTQQWQSNNVAPPHLFATAAASSGF
PQQILRHQNWPHKNGFHYSLTRPSN*
>StAP2-02
MWNLNDSPDHRRDYEGKGVERSISNSSSSAVEEL
YGSEEEDEFLEQNGLKGKKKKSIPSK
IFGFSMIAPPNNNNDDNLSSESEPPVTRQFFPVD
ESEIGSGNFNDRSCRFPRAHWAGVKF
YQPEPSTAVLGKGSELSQVQVQVQVQVQPMKKSR
RGPRSRSSQYRGVTFYRRTGRWESHI
WDCGKQVYLGGFDTAHAAARAYDRAAIKFRGVEA
DINFNLEDYEEDLKQMKNLTKEEFVH
VLRRQSTGFPRGSSKYRGVTLHKCGRWEARMGQL
LGKKYVYLGLFDTENEAARAYDKAAI
NCNGKDAVTNFDPCIYENELNSSESSNKAADHSL
DLSLGSSTSKQKSREMEDTNKDQNYS
SMQFDVDWRHQGSRLKQHSPIDMEYGRRRNVYNE
TDTLQLLSQTHLHSPGSFKTNSNEMQ
RFGQYMRAGDSQMIPPQFTSSNYQVQFPSSNNRR
DTQQWQSNNVAPPHLFATAAASSGFP
QQILRHQNWPHKNGFHYSLTRPSN*
>StAP2-03
MWDLNESEEGCSTPIEFECDDEKGKRVGSVSNSS
SSAVAVDDISEEELDGERGKKKRGTP
PSKIFGFSMVGPGNGDEEQPVTRQFFPVDESEMG
GVAAAENGSPNFPRAHWVGVKFYQTE
TLGNTGLAKPVDMVQQQQQQPIKKSRRGPRSRSS
QYRGVTFYRRTGRWESHIWDCGKQVY
LGGFDTAHAAARAYDRAAIKFRGVEADINFSLED
YESDLKQMTSLTKEEFVHVLRRQSTG
FPRGSSKYRGVTLHKCGRWEARMGQFLGKKYVYL
GLFDTEIEAARAYDKAAIKCNGKDAV
TNFDPSIYEDELNSIESSDNAADHNLDLSLGGSS
SKQGSRELGDNRGQNSSSMQLDIDWQ
RHGLRPEKQSSALIDARRRENGYNETETLQLLSQ
THLHSPVSLKHNNSQVQRFSQFMRPG
DQSHMIQMFPPQFGSSNYQIQFPSGSNGGRIGGT
NVGDLSLAATSNGSSQWQSNFPPQIF
AAAAASSGFPQQIVRPQNWSSENGFHHSLRRPS*
>StAP2-04
MWDLNESEEGCSTPIEFECDDEKGKRVGSVSNSS
SSAVAVDDISEEELDGERGKKKRGTP
PSKIFGFSMVGPGNGDEEQPVTRQFFPVDESEMG
GVAAAENGSPNFPRAHWVGVKFYQTE
TLGNTGLAKPVDMVQQQQQQPIKKSRRGPRSRSS
QYRGVTFYRRTGRWESHIWDCGKQVY
LGGFDTAHAAARAYDRAAIKFRGVEADINFSLED
YESDLKQMTSLTKEEFVHVLRRQSTG
FPRGSSKYRGVTLHKCGRWEARMGQFLGKKYVYL
GLFDTEIEAARAYDKAAIKCNGKDAV
TNFDPSIYEDELNSIESSDNAADHNLDLSLGGSS
SKQGSRELGDNRGQNSSSMQLDIDWQ
RHGLRPEVYLHDQLLYLVFVVNFFFFLCNMF*
>StAP2-05
MIVALDELNHHCGISVYLGGYDKEEKAARAYDLA
ALKYWGPTTTTNFPVSNYERELEEMK
NMTRQEFVASLRRKSSGFSRGASIYRGVTRHHQH
GRWQARIGRVAGNKDLYLGTFSTQEE
AAEAYDIAAIKFRGLNAVTNFDISRYDVKSIASS
NLPVGGMSNKSKSSSDESTKNVENRP
NDISSAPNFPSQNPSMVNLGFGLPIKQDASDFWS
SLGYNNDQNKNGSLFQGTTMNAQGTT
FFSNEVNNSNVMFNEQGYYNHDQQQQSGGSSSEI
SMALNSNIATTTTTIDSSSFGNWMTP
SLHSFQSSANNNLAPFHTPIFGME*
>StAP2-06
MKKSPSFSCSSSSSSSSCIEQIHEETEKLNLKPK
PKPRLKRASRAKKIVNADSPSNNSST
VRRSSIYRGVTRHRWTGRYEAHLWDKSTWNSIQN
KKGRQRAYDSEEAAARTYDLAALKYW
GPTTILNFPVERYSQEFEEMQRLTKEEYLASLRR
MSSGFSRGVSKYRGVARHHHNGRWEA
RIGRVYGNKYLYLGTYSTQEEAAAAYDMAAIEYR
GPNAVTNFDISRYADHLKKLREPNLL
TKEENTESSAEVQSNEVIEQYQPVQQEENQLDYQ
LAELAAEPIVVPKLEFTPALDSDEVT
QPKVLKLEFAADIMKAKDHEEEEDYPWMNMYLD
NTFDSLPVSDFSLDKPADLMDLFNDNS
FDNNIDFNFYEQSSENEFNLNVFSDSMIIDGIEA
DNEEVRNNLSISPTSSSSMSRTTSIS
NDMKDDGGLASSTM*
>StAP2-07
MASKDEHKENEAVEEPSEIPRDEKSINPVGESSV
VLPRVVRIYMPDSDATDSSSDEEEKL

QGEKSKRHERTCIKEIIIENGKTRVICKKTSKEK
KDIKLFQENAKKYKGVRQRRWGKWAA
EIRDTKNKTRLWLGTFDTAEEAALAYDKAAIEIR
AMTSKEDLQVKEGVIEPSAIDVVEKS
SNQVGEISQVFPRLVRIHVPDHDTTDSSSDEEEK
DEEGKSEGHERTCVKEIIIENGKTKV
ISKKKSKEKKDLNLMQEDVNKYRGVRRRKWGKWA
AEIRDTRKKTRLWLGTFDTALEAALA
YDKAAIEIRGANALTNILKPPAKKNNPIKINFIV
PPLKRQKADP*
>StAP2-08
MKTTITTNTNTTKSSSTQPKKSSITRKFVGVRQR
PSGRWVAEIKDSSQRVRLWLGTYDTA
EEAAHAYDEAARALRGENARTNFASTAPNSDMDQ
SHSSHDTKHGLSFSSIKEKLSKNLQS
IMARNSENKSSSSSKSITRVSDHFTFARIFHFKN
NYDQQYQNHHRHVDMNKVVQPSIRVL
PQDVTDNNNNNDLCWENSSSVSDCSSEWAAFRQL
GLDSDNYGSDGSDQYFVGSDPLMAGW
ISSPDIMSNSTNEGSSRSKRFKVSSSVMVPPTFS
EKIMYHFN*
>StAP2-09
MAVLDQTSNMMLPMDYTRKKKSRSRKDAPKNVA
DTLAKWKEVNEKLDACDDDGRKPVRKV
PAKGSKKGCMKGKGGPDNGRCKYRGVRQRTWGK
WVAEIREPHRGRRLWLGTFDTAIEAAL
AYDEAARAMYGPCARLNLLDYYPSSKESSKDDSS
LPTESHSDSNTASSFSEVCPADDMMR
GRANVPEIRYEDGASEIKFDGTRTGSNKVGTPLS
SLREEAKDEAKEVLDKSETFTPLRSL
REEAEDKAKEVLDKSETFEIKEEPVACSYDSWDI
GQEDLGNLCLDDEMFDVNELLGMMDS
TPLDASAPSQDVGFVPPKQEQYAYDPSYRLQNAA
YDNNQLSNPSYQLENADDQFSNPYQL
DNADVNTLEGLQQMEQQPPIEIDYDFDFLRPGRQ
EDFNFCLDDLGVLDF*
>StAP2-10
MAMKEKVTNGVKKKEVHYRGVRKRPWGRYAAEI
RDPNKKTRVWLGTFDTAEEAAKAYDAA
AREFRGLKAKTNFPFPSEDQSPNDSSGSPNESFS
GENGVHALPELDHTRRLEEGGGGYTE
VGLVRNGFPIFRQQPAMMSNRSSPAESSIGETVV
QSPFELNLTRRLEGGGNGGRSAEVGL
LRNGFPMFLHQPASHAELDLTRNLGDHVRNGFPI
SHRQPAVAVLPNGQPILLYDGNLVRT
GLVNRPQPYRFEPVATAAEFNGVTGRVVDSESDS
SSVVGLGRRELNLDLNLAPPMEV*
>StAP2-12
MVDSNGNNAVYLGAYDDEGTAAHAYDLAALKYW
GSETILNFPVSTYEKEIIAMENQSREE
YIGSLRRKSSGFSRGVSKYRGVARHHHNGRWEAR
IGRVFGNKYLYLGTYATQEEAATAYD
MAAIEYRGLNAVTNFDLSRYIKWLKPNKNNQSNP
NVETNMISTPNHTTTITTTTTTSSSM
GLVYPDLQLTTNTPISDVSTSQSRPSSATSALGL
LLQSTKFKEMMEMTLAVECPPAPIEL
PDPDPPQSSSFPEDIKTYFDSQEFGNFGHDQGED
LVFNEINSLMQPLLQFDYDA*
>StAP2-11
MAIMDQAANMATLPLDYSGKRKSRSRRDRTKNVE
ERLAKWKEYNEKLDSVDDEGKPVRKV
PAKGSKKGCMRGKGGPENSRCKYRGVRQRIWGK
WVAEIREPKRGSRLWLGTFGTAIEAAL
AYDEAARAMYGPCARLNLPNYPSRDDSDSASWAT
TSASDCTVASGGFGGEVCPGEGATTA
AGTPLSSVKDEGKDETAGLPGEMEIVEPTSIDQD
TLKSGWDCLDNLNLDEMFDVDELLAM
LDSTPVSTNDFGSDGKQYAYDNNLLSNSPCQPLG
DPQEMGQEAPMTFDYAFDFLKPGRQE
DLNFNSDDLTFMDLDSELVV*
>StAP2-13
METVTASGKLKELSFGQIENSGTSNSSIVNVETS
STAGDDEFISCSDQRTDGYAFEILRA
DYEGNGFVTKELFPLTGGESAAPPSQQQQWLDLS
GNYSDVPMEQRIVLGPPQLRQQVKKS
RRGPRSRSSQYRGVTFYRRTGRWESHIWDCGKQV
YLGGFDTAHAAARAYDRAAIKFRGLD
ADINFNVSDYHDDLKQMGNFSKEEFVHVLRRQST
GFSRGSSKYRGVTLHKCGRWEARMGQ
FLGKKYIYLGLFDSEIEAARAYDKAAIKCNGREA
VTNFELSTYEGELSTEADNGGADHNL
DLNLGIAPSVFADDQGGNTCQMGNYVVQCVSNGL
PGYGGAMNIPTTTMGGKLLHGHQLLT
DRPLLWNGVNTSIFPTLKGTALGKGIEVDSSPNW
TRQDQHPYGGSPSVPLFSTAASSGFA
NTTTAAAQLPCFSNGKLPYQHSPSLANMNFAHYY
CRS*
>StAP2-14
MEQDPSNGRRTNGGGGGGGEIKYRGVRRRPWGKF
AAEIRDSARQGARVWLGTFNTAEEAA
RAYDRAAYSMRGHLAILNFPEEYNLPTSSSHFYS
AGSYSSPSMASSSSSSSRQVLEFEYL
DDKLLEELLDCDEEPNKRK*
>StAP2-15
MNEDNSNTPKKKESCVEVKYRGVRKRPWGKYAAE
IRDTNKNGARVWLGTYATAEEAARAY
DKAAFEMRGRFAVLNFPHEYPSDHHSSSYSMSSS
STSCSSVHNGNEMIKEVIELEYLDDK
LLEELLGMESDKSHEYNK*
>StAP2-16
MWNLNDSPDQIMEDESDEGIPVRSVSNSSSSAVL
VAEDGNSSEEDGERGKKKKSSTPGKL
FGFSVNDDNLEPPIVTRQFFPVDNESTNFPRAQW
AGIKFCESEPPLVTGLVGNPIEVSQQ
QPIKKSRRGPRSRSSQYRGVTFYRRTGRWESHIW
DCGKQVYLGGFDTAHAAARAYDRAAI
KFRGAEADINFTLEDYEGDLKQMSNLTKEEFVHV
LRRQSTGFPRGSSKYRGVTLHKCGRW
EARMGQFLGKKYVYLGLFDTEVEAARSYDKAAIK
CNGKDAVTNFDPSIYENEVNSTECSD
NAADHNLDLSLGGSSQEMGDNRGQNSSSNLQLDV
HWGHQGSSRLNKPRSQINFDGQKRSG
YNETETLQLLSQTHLNSSGRVQPNINEMQKFGEF
IRDGETHMVQAYPRQFNSYQVQTSSN
NNVGQIGATSNSQEWHPNYEIITTAAASSGFPQQ
IIRTQNWSQNNGFHHNFMRPS*
>StAP2-17
MDEVKYRGVRKRPWGKYAAEIRDTNRQGGVRVW
LGTFGSAEEAARAYDKAAYNMRGHLAI

LNFPEEYNLPRSSSHFYDNSQSMASSSSSNVHAR
NDTGQVLELEYLDDNVLEELLNCDQH
ANKRK*
>StAP2-18
MIKNWLRNNPNSTTTTTTTPSLLQDDKKEGDVVG
VAGGGSSGTNAQSLSLSMSTGGDGGG
ENSCSSENKQEIVNGNNGSTIDGQISGAIVAPKK
SIDTFGQRTSIYRGVTRHRWTGRYEA
HLWDNSCRREGQTRKGRQGGYDKEEKAARAYDLA
ALKYWGPTTTTNFPISNYEKELEEMK
NMTRQEYVASLRRKSSGFSRGASIYRGVTRHHQH
GRWQARIGRVAGNKDLYLGTFSTQEE
AAGAYDIAAIKFRGLNAVTNFEINRYDVKSILES
STLPIGGAAKRLKDVEQAEIALNYQR
ASQENIGTTHLMNGSMSTYGAGHCWPNITLQQAQ
PLSAMHYNNPYNSQQRLWCKQEVQDH
DLTSQNFQDFQLGNAHNYLMGLDNSSSSSMEQNN
VTYHQGVGYASNNNNHHGGGGFVLPQ
LNGSSYGGENDNEVKQLQLGTHENIFGNYYHSQP
TTSNDSVKVSNTLYDHQESACNNNNW
MSTDVPMALASRTSTMAVCHSAAPTFTMWNDS*
>StAP2-19
MMLDLNLSAIYDGKHSLADESGSSNSSARNVEAS
SSAGDDETCSTRAAAGSGDMFAFNFD
ILKVGGGGGGENSRSCSNNDDEEGYVITRSEFVT
QQLFPVDNNAELNRTHTSRTTDWVDP
SVDPPNTVSFREVEQMGRLQQQQQQQQQPVKKSR
RGPRSRSSQYRGVTFYRRTGRWESHI
WDCGKQVYLGGFDTAHAAARAYDRAAIKFRGVD
ADINFNLSDYEEDMKQMKNLSKEEFVH
VLRRQSTGFSRGSSKYRGVTLHKCGRWEARMGQF
LGKKYIYLGLFDSEVEAARAYDKAAI
KCNGREAVTNFEPSTYEGETNSNPQSEGSQQDLD
LNLGIATSSPKEIERSSSFQYHPYDM
QEATKSQMDKSGSPIVGSSHYKALPAVTSEQAHL
WNGVFSNLIPNYEERVSGKRIDVGTS
HGHPNWAMQMHSQVGTTPMSMFSAAASSGFSTP
ATTASASQMSGPNNANTLNLSFASYST
PSMNTTPYYYQIRPQLPPP*
>StAP2-20
MNSNNWLSFPLSPTHSSLPPHLQNDQSHHFSLGL
VNETIDNPFQNQEWNLMNTQGSNEVP
KVADFLGMNKSENQSELIPYNEIQANDSDYLFQN
NHLMPSMQNALAPPPTSNYDLQENAC
NIQSLTLSMGSGKGSTSETSASPSATAATASASA
ENSNNTSIVEAAPRRTLDTFGQRTSI
YRGVTRHRWTGRYEAHLWDNSCRREGQSRKGRQ
GGYDKEEKAARAYDLAALKYWGTSTTT
NFPISNYEKECEDMKHMTRQEFVAAIRRKSSGFS
RGASMYRGVTRHHQHGRWQARIGRVA
GNKDLYLGTFTTEEEAAEAYDIAAIKFRGLNAVT
NFDMNRYDVKAILESNTLPIGGGAAK
RLKEAQALESSRKRDQEMMALNSSFQYGNSNPLQ
AYPLMQQQPSFDNSQPLLTLQNHDIS
QYNIQDSSSHFHQSYLQTQLQLQNNSHQVLYNNY
LQNNQVFLHGLMNNNEGSSSGSYSTG
GYFGNSPGLGGMSSNSTSGNNGGGGGGSGSVAHE
EVALVKVDYDNIPSYNGWSGESSVQG
SNPGVFSMWNE*
>StAP2-21
MAGSRFLAIIQSTVYLGGYDKEDKAARAYDLAAL
KYWGPTATTNFPVTNYTKELEEMKHM
TKQEFIASLRRKSSGFSRGASIYRGVTRHHQQGR
WQARIGRVAGNKDLYLGTFATEEEAA
EAYDIAAIKFRGVNAVTNFEMNRYDVEAIMQSSL
PVGGTAKRLKLSLESEQKSSSSNNNH
QQQQQLTQCNSTNNINFGAMPPVSAIPCGVPFDN
TTPAFYHHNFFQYLHPGNASVPDASN
AAAMPLLPSAADLFIWPHQSY*
>StAP2-22
MDIKEGLSSKKLVVNIKQGMKLESDECNTSTNCN
PKAKISLPGIRRQKNGRYGVELTDPI
RHKRVWVGTFDTVEEASQAYSSKKSEFENEKLRQ
QGNKKSKRIRKTIGMVEVHKRGKYNS
EKSEELVNVKQENENVNREIFQTESAGGSEIDFP
ISNSSNGGTEKRIDLHEIGAVEEAFR
IAKEGSRVFKMANVCGTKSSDDCNNTTSCNPEVK
VSLIGTRKQKNGRYAAVITDPFTHKR
VRLGTFDTVEEASQAYFSKKSEFEKLSQQGNKQN
ILKKKNCDQIQQLESSSVVVASLESG
SGSGRARRIDSHKKKTHLIGAHKSTKKGRYQSEI
KNPITKKRIWLGTFGSAEEASHAYQS
KKLEFQKLVEEKKKQCTYKQLKSEKLVNLKEGHE
NVKCEPFLPESAPDIDIPFSNSSNGG
TVQSIDPDKIGNQSKKFDLELESNMPIGSSAGEK
QEGEEDDEDLWTGQWVQLPGDDRAVM
FSLKLGLPIIDNYGSLLGEFSTLDDLSICKTEEG
NNYISHSI*
>StAP2-23
MNAISVKCEKLSPGNRRSLCMVDYDIPVNNKCVK
RRRKSPVAAVIVNDNNNNGQKLDSSK
VDQSSSIVQTRTVKRSSRFRGVSRHRWTGRYEAH
LWDKASWNVTQKKKGKQVYLGAYDEE
ESAARAYDLAAIKYWGTTTFTNFPISDYEKEIEI
MQNMTKEEHLASLRRRSSGFARGVSK
YRGVARHHHNNRWEARIGRVFGNKYLYLGTYNTQ
EEAARAYDIAAIEYRGTNAMTNFDMN
TYIRWLKPDANSLMQFQDLTSEDQPIQTSNTNDM
INPLHDSSYSFNHNHFTTTTNTEKLY
AIPRSQEPIERKMPLSYCRKSSSPTALSLLLRSS
MFQELVEKNSSTDSEESEENNMKHGA
QSNRENDFLFYTNDKLPKLESLENGTTSSSALCD
REGQQYFGIDES*
>StAP2-24
MIQTRITLFSPNPIPNLAVPAASRKEREKEKERT
KLRERHRRAITSRMLAGLRQYGNFPL
PVRADMNDVLAALARQAGWTVEPDGTTFRQSPAP
APTNASNMNDDPNSALSILKRTCPAE
KRNNNNNVPSVEKSYRGVRNRPWGRWSAEIRDRI
GRCRHWLGTFDTPEEAARAYDAAARW
LRGIAAVVGEHVLFGYNNNENKNDDPNSAVSILK
RTCPAEKRNNVPSVEKSYRGVRNRPW
GRWSAEIRDRIGRCRHWLGTFDTPEEAARAYDAA
ARWLRGSKARTNFQIPPIVPLPTSPT
STSSSSNSSREMKMKMKMKNGGAVANNQRKCSVG
TYQVMSVESPVSGSSFRNCSTRASVD
GQPSVPRINESLSPASFDSIVVTESDTKVDKFTS
TSPMNSTGCLEAGQLMQELHCGEHGS
SFAETQYVPVFVMLSSGVINNFCQLMDPDGVKQE
LQQLKSLKIDGVVVNCWWGIVESWVP

QKYEWSGYRELFKIIRDFNMKLQVVMAFHENGGS
DTSGMFISLPQWVLEIGKDNQDIFFT
DRQGRRNTECLSWGIDKERVLRGRTAIEVYFDLM
RSFRTEFDDLFTDGLISAVEIGLGAS
GELKYPSFSERMGWRYPGIGEFQCYDKYSLQNLR
KAATSRGHSFWAKGPDNAGYYNSKPH
ETGFFCERGDYDSYYGRFFLHWYRQVLIDHADNV
LTLATLAFEGIQIVVKIPAIYWWYRT
SSHAAELTAGYYNPTNQDGYSPVFEVLKKHSVTI
KFVSSGLQVPETDDALADPEGLSWQI
MNSAWDKEINVAGNNAFPCYDREGFMRLVETAKP
RNDPDRHRFSFFAFQQPSPLVQSAIC
FSELDYFIKCMHGIVF*
>StAP2-25
MSMETDEEGLLADEKEEEESSSKKQKITQIEQSE
IFVDKVKSTIPLQHTNHKAQSTADRS
KQETDGKMPLFVNEWQMGEKGAFLGTEEFKGFSS
NMHKFQEKEPFICMESCSKETPCLLS
NVHGGIESSNKAKINLIGVRRQKNGRYGAVITDK
IRHKQVWLGTFDTVEEASQAYLSKKS
EFEKFKQQGNKDNKPKENLDQIQQTSLFVGNDQT
LDTASVSRINPHETIHVVEVQNKNWL
GKEPESAKETTCLMVNVHGIESYDECNTTTSCNP
KVKRSLLGVRKQKTGRYGAVITDQIR
HKQVWLGTFDTIEEASQAYFSKKSEFEKLKQQGN
KDTENCDQIQQPESPVVLSSSSVAND
QTLDTASVGCRNRRIDSHETAPHIVEVHKNKMSG
EVPESSKETQRLMASVVDNTESSDEC
NTSTSNLKAKISLLGVRRQKNGRYGAVITDTTKH
KQVWLGTFDTVEEASQAYFSKKSELE
NEKLRQQRNKENRPKKNCDQIQQPESPVVQASLS
VADVQTLGTASVGMRNERTVFHGTAH
VVGVHKSKTSEKEPEFTKETSCLMDSVHATESSD
QCNTSTSNSTAKISLIGVRRQKNGRY
GAVITDTTRHKKVWLGTFDTVDEASQAYFSKKSE
LENEKLNQQGNNETRDQIQQPESLVA
SLSMDNDQTLNAASGGRRNKRIDSHKATPDIVGV
HKNKTSGKEFESSRETARLMDNVHGT
ESCDECNTTTSCDPKAKISLIGIRRQKKGRYGAV
ITDRIKHKRVWLGTFDTVEEASQAYL
SKKSELKRLGQQGDEENKLKDCAQVQQPESPVVP
SLSVANHDQTLNTARMRRNERIDSHK
TITRFFRVHKRKDSGKYTSEIRNPISKKRIWLGT
FGTAEEASQAYQSKKLEFERLVHAQQ
QCGNEQTHSTQDEKSEKLVNVESGHENVNRELES
AGGSEINVTISNSSNGGTEERIDSHE
IGTAEEAFHAYQSKKFDMPTDSSAGEKQEGQEDD
EDLEMGEWVQLPGNRAVKFSLKLGLP
IIDNYGSLLGEFSTLDDLSICKTDYDNETYMCLS*
>StAP2-26
MDSKKLVVNVKQGEKLEYDECNTSKNCNPKDKIS
LIGIRRQKNGRYAAVITDRIRHKKVW
LGTFDTVEEASQAYFSKKSEFENEKLSNQGNKKS
KRKRETRSIVEVYKRGKYNSEKSEEL
VNVKQGNENVNCEVFQIESAGESEIDVVMSNSFN
GGTEQRIDSHEIGTVEEAFRIAKEGS
KVFILGLMANVCGIKSFDDCSNTTSCNPKVKVSL
IGTRRQKNGRYTAVITDPFTHKRVRL
GTFDTVEEASQAYFSKKSEFEKLSQQGNKQNIPK
KNCDQIQQLESSSVVASLDTAASESG
IAKRIDSQGEEPKSSKATSCVKANAHSAESSDEC
SITTSCDHKPRISLTGIRRRKSGRYA
AVITDPIKHKEVCLGTFDSIEEASHAYLSKKSEF
EKLRQQGDKKNKPKKNCDQIQQPESQ
SVVASLDTTTRDSGSAKRIDSHKTTTHLIGAYRS
KTAGRYRSEITNPITKKKIWLGTFDS
AEEASHAYQSKKLEFQKLVEAKQQQCTNKQRKSE
KLVSVRQGHENVNCEPFHLESASGSP
DIDVLFSNSSNEGTVRSIDSYQSNKFDLLSSKEV
ELQSNMLTESSAGKKQEGQEDDKDLW
KGEWMQLPGDNRAVMFSLKLGLPIIDNYGYLLGE
FSTLDDLSISVTEDDNH*
>StAP2-27
MKAMNDDSRSNNNNNHNSAAANSNWLGFSLSPH
MKMEVTNSSETQHQHQFAQSFYLSSSP
PPPMNVSTTSALCYENNPFHSSLSVMPLKSDGSL
CIMEALSRSHADAMVQSSSPKLEDFL
GGASMGNSQYGSHDREAMALSLDSMYYHQNEEEI
QVHSHHPYYSPMHCHGMYHQESLLEE
TKPTQISNSDAQMTGNELKSWGHYAEQHAIDQHI
NATCSMVAAAAAAAGGGGGTSGCNDL
QSLSLSMNPGSQSSCVTPRQISPTGLECVAVESK
KRASGKVAQKQPVHRKSIDTFGQRTS
QYRGVTRHRWTGRYEAHLWDNSCKKEGQTRKGR
QVYLGGYDMEEKAARAYDQAALKYWGP
STHINFPLENYQKELDDMKNMTRQEYVAHLRRKS
SGFSRGASIYRGVTRHHQHGRWQARI
GRVAGNKDLYLGTFSTQEEAAEAYDVAAIKFRGV
NAVTNFDISRYDVEKIMASNTLPAGE
LARRTKEREPIEYNNNNNNNNIGGVGGHKNEECV
DNNNNGNITDWKMVLYQTSNPSLGLN
YRNPTSFSMALQDLIGIDSMTNSNNHHATILDHE
QNKIGNHFSNASSLVTSLGSSREPSP
DKSATAASLVFAKPTKFAVPTATSVNACIPSAQL
RPIPVSMAHLPVFAALNDA*
>StAP2-28
MENSGTSNSSVVNMETSSTAGDDEYFSSDHRHRD
GFTFAILKADKEANNEYEGSGDRGGF
VIKELFPLVNGEVGLDLPANCGDLTEQRIGKTPQ
QRLQAKKSRRGPRSRSSQYRGVTFYR
RTGRWESHIWDCGKQVYLGGFDTAHAAARAYDR
AAIKFRGLDADINFNISDYEEDLKQMK
NFSKEEFVHILRRQSNGFSRGSSKYRGVTLHKCG
RWEARMGQLLGKKYIYLGLFDNEIDA
ARAYDKAAIKCNGREAVTNFEPNTYEGALTLEAD
IGGRDHNLDLNLGISLSSYAENQHGN
TSQIGNFQCRHGSNGLPEHHGEVRTPASTTLRSK
LLLHGQHMLTQHPLHWNGSNDNLFPT
FKVSAIEKGLEVDASLKWMGHDQNTYDGSPTLPL
FSTAASSGFGNSANTAPSAATHQLHF
GSGALPFPHSPSLTNMNLSHHYFRS*